



USER MANUAL

10GATS-8-L23

L3 Managed Ethernet Switch

CLI User Manual

Ver 1.01

PoE Texas

Contents

Notices	1
Copyright.....	1
About This Manual	1
Intended Audience	1
Conventions	1
Chapter 1 CLI Command Line Introduction	3
1.1 Accessing the CLI of the Switch	3
1.2 Introduction to CLI Mode	5
1.3 Command Syntax Introduction	8
1.4 Command Line Shortcut	11
1.5 History Command	12
Chapter 2 System Management Configuration	13
2.1 System Security Configuration	13
2.2 System Maintenance and Commissioning	17
2.3 Configuration File Management	20
2.4 Software Version Upgrade	24
Chapter 3 Configure Port	28
3.1 General Configuration of the Port.....	28
3.2 Configure MIRROR.....	29
3.3 Configure STORM-CONTROL	31
3.4 Configure STORM-CONSTRAIN.....	32
3.5 Configure Flow-Control	34
3.6 Configure Port Bandwidth	35
3.7 Configure TRUNK	36
3.8 Configure Jumbo Frame.....	39
3.9 Configuring Redundant Port.....	40
3.10 DLDAP Configuration	41
3.11 Configure LLDP	42
Chapter 4 Configuring Port-Based Security	44
4.1 Introduction	44
4.2 MAC Binding Configuration	44
4.3 MAC Filtering Configuration	45
4.4 Port Learning Limit Configuration	46
4.5 Protection Port Configuration	47

Chapter 5 Configure Port IP and MAC Bindings	49
5.1 Introduction	49
5.2 IP and MAC Binding Configuration.....	50
5.3 Configuration Examples.....	50
5.4 Configuration Troubleshooting	52
Chapter 6 Port Loop Detection	53
6.1 Introduction	53
6.2 Protocol Principle	53
6.3 Configuration Introduction.....	54
Chapter 7 Configuring VLAN	56
7.1 Introduction to VLAN.....	56
7.2 VLAN Configuration	60
7.3 VLAN Configuration Examples.....	63
7.4 MAC, IP Subnet, and Protocol VLANs.....	67
7.5 Voice VLAN.....	68
7.6 VLAN Mapping	70
7.7 QinQ	71
Chapter 8 Configuring QoS	75
8.1 QoS Introduction	75
8.2 QoS Configuration	77
8.3 QoS Configuration Example.....	79
8.4 Policy QoS Configuration Example.....	79
Chapter 9 Configuring MSTP	81
9.1 MSTP Introduction.....	81
9.2 MSTP Configuration	86
9.3 MSTP Configuration Example	95
Chapter 10 Configuring EAPS	98
10.1 EAPS Introduction.....	98
10.2 EAPS Basic Concepts	98
10.3 EAPS Protocol Introduction	98
10.4 EAPS Configuration	100
10.5 Restrictions	100
10.6 Brief Introduction to EAPS Commands	101
10.7 Single Ring Configuration Example	102
10.8 Cross Ring Data Forwarding Configuration Example	107
Chapter 11 Configuring ERPS	111

11.1 ERPS Overview	111
11.2 Introduction to ERPS Technology	111
11.3 ERPS Operating Principle	112
11.4 Technical Features of ERPS	114
11.5 ERPS Protocol Commands	116
11.6 ERPS Typical Application	117
Chapter 12 Configuring AAA	136
12.1 802.1x Introduction	136
12.2 RADIUS Introduction	141
12.3 Configuring 802.1x	144
12.4 Configure RADIUS	148
12.5 Configuration Examples	151
12.6 TACACS+ Introduction	151
Chapter 13 GMRP Configuration	153
13.1 Introduction to GMRP	153
13.2 Configure GMRP	153
13.3 GMRP Typical Configuration Example	154
Chapter 14 Configuring IGMP SNOOPING	156
14.1 IGMP SNOOPING Introduction	156
14.2 IGMP SNOOPING Configuration	160
14.3 IGMP SNOOPING Configuration Example	164
Chapter 15 MVR Configuration	165
15.1 Introduction to MVR	165
15.2 Configure MVR	165
15.3 MVR Configuration Example	166
Chapter 16 Configuring DHCP SNOOPING	168
16.1 Introduction to DHCP SNOOPING	168
16.2 DHCP SNOOPING Configuration	169
16.3 DHCP SNOOPING Configuration Example	171
16.4 DHCP SNOOPING Configuration Troubleshooting	173
Chapter 17 DHCP CLIENT Configuration	174
17.1 Introduction to DHCP CLIENT	174
17.2 DHCP CLIENT Configuration	174
Chapter 18 Configuring DHCP RELAY	175
18.1 Introduction to DHCP RELAY	175

18.2 DHCP RELAY Configuration.....	176
18.3 DHCP RELAY Configuration Example.....	176
Chapter 19 Configuring DHCP SERVER	178
19.1 Introduction to DHCP SERVER	178
19.2 DHCP SERVER Configuration.....	179
19.3 DHCP SERVER Configuration Example.....	182
Chapter 20 Configuring ACL	185
20.1 ACL Resource Library Introduction.....	185
20.2 ACL Filtering Introduction.....	186
20.3 ACL Resource Pool Configuration.....	188
20.4 ACL Time-Based ACL	190
20.5 ACL Filtering Configuration.....	192
20.6 ACL Configuration Example	193
20.7 ACL Configuration Troubleshooting	194
Chapter 21 Basic TCP/IP Configuration	195
21.1 Configure VLAN Interface	195
21.2 Configure ARP	196
21.3 Configure Static Route.....	198
21.4 TCP/IP Basic Configuration Example	200
Chapter 22 Configuring SNMP	202
22.1 Introduction to SNMP	202
22.2 SNMP Configuration	203
22.3 SNMP Configuration Example	204
Chapter 23 RMON Configuration	205
23.1 RMON Introduction	205
23.2 RMON Configuration	205
23.3 RMON Configuration Example	207
Chapter 24 Cluster Configuration	208
24.1 Introduction to Cluster Management	208
24.2 Introduction to Cluster Configuration.....	213
24.3 Configure the Management Device	214
24.4 Configure the Member Device	219
24.5 Configure Access to Cluster Members	219
24.6 Cluster Management Display and Maintenance	220
24.7 Cluster Management Configuration Example.....	220

Chapter 25 SNTP Configuration	224
25.1 Introduction to SNTP	224
25.2 Configuring SNTP	224
25.3 Information Display of SNTP	226
Chapter 26 Configuring PTP	227
26.1 Introducing PTP	227
26.2 Configuring PTP	227
Chapter 27 Configuring IGMP	229
27.1 IGMP Introduction	229
27.2 IGMP Configuration	230
27.3 IGMP Configuration	233
Chapter 28 Configuring PIM-SM	236
28.1 PIM-SM Introduction	236
28.2 PIM-SM Configuration	237
28.2.2 Configure Multicast Routing Table Capacity	238
28.3 PIM-SM Configuration Example	247
Chapter 29 Configuring RIP	250
29.1 RIP Introduction	250
29.2 RIP Configuration	251
29.3 Sample RIP Configuration	259
Chapter 30 Configuring RIPng	262
30.1 RIPng Introduction	262
30.2 RIPng Configuration	262
30.3 RIPng Configuration Example	267
Chapter 31 Configuring OSPF	269
31.1 OSPF Introduction	269
31.2 OSPF Configuration	270
31.2 Sample OSPF Configuration	284
Chapter 32 Configuring OSPFv3	286
32.1 OSPFv3 Introduction	286
32.2 OSPFv3 Configuration	288
32.3 OSPFv3 Configuration Example	296
Chapter 33 Configuring BGP	299
33.1 BGP Introduction	299
33.2 BGP Configuration	301

33.3 BGP Configuration Example	306
Chapter 34 Configuring VRRP	308
34.1 VRRP Introduction	308
34.2 VRRP Configuration	316
34.3 VRRP Configuration Example	320
Chapter 35 Configuring VLLP	322
35.1 Introduction to VLLP	322
35.2 VLLP Configuration	325
35.3 VLLP Configuration Example	327
Chapter 36 Configuring Policy Routing	330
36.1 Introduction to Policy Routing	330
36.2 Policy Routing Configuration	330
36.3 Policy Routing Configuration Example	331
Chapter 37 Configuring the System Log	333
37.1 System Log Introduction	333
37.2 Syslog Configuration	336
37.3 Configuring SYSLOG	338
Chapter 38 IPv6 Basic Configuration	340
38.1 Introduction to IPv6	340
38.2 Introduction to Basic IPv6 Configuration Tasks	349
38.3 Configure the Basic Functions of IPv6	349
38.4 Configure the IPv6 Neighbor Discovery Protocol	350
38.5 IPv6 Static Route Configuration	353
38.6 IPv6 Display and Maintenance	353
Chapter 39 MLD SNOOPING Configuration	354
39.1 Introduction to MLD SNOOPING	354
39.2 MLD SNOOPING Configuration	358
39.2.1 MLD SNOOPING Default Configuration	358
39.3 MLD SNOOPING Configuration Example	361
Chapter 40 PoE Configuration	362
40.1 Introduction to POE	362
40.2 Configure POE	362

Notices

Copyright

© PoE Texas. All rights reserved.

No part of this manual may be extracted, reproduced, or distributed in any form by any organization or individual without the prior written permission of PoE Texas.

About This Manual

This manual is the command-line reference for the PoE Texas 10GATS-8-L23 L3 Managed Ethernet Switch. It describes how to reach the command-line interface, how the command modes are organized, and the commands available for every feature the switch supports.

Each feature chapter follows the same shape: an introduction to the feature, the commands that configure it — with the command syntax, what it does, and the mode it is entered from — and, where the source provides one, a worked configuration example.

For the equivalent configuration through the browser, see the 10GATS-8-L23 Web Interface User Manual.

Intended Audience

This manual is written for:

- Network planners and designers
- On-site technical support and maintenance personnel
- Network administrators responsible for network configuration and maintenance

Conventions

Command syntax in this manual uses the following conventions.

Convention	Meaning
Plain text	A keyword you enter exactly as shown.
<angle brackets>	A parameter you supply a value for, such as <if-name>.
[square brackets]	An optional keyword or parameter.
{ braces }	A required choice. Alternatives are separated by a vertical bar, as in {normal privilege}.
vertical bar	Separates the alternatives in a choice. A space is required before and after the bar.
asterisk *	The preceding keyword, parameter, or group may be entered more than once.

The rest of the manual uses these conventions:

Convention	Meaning
Monospaced text	Text you type at the prompt, or output the switch returns.
Switch#	The CLI prompt. The part before the mode character is the host name, which defaults to Switch.
Note	Supplementary information that helps you use a feature correctly.
Caution	Information you must read to avoid losing configuration, losing management access, or damaging the switch.

NOTE

Command examples in this manual use the default host name, Switch. If you change the host name, your prompt will differ from the examples.

Chapter 1 CLI Command Line Introduction

This chapter describes the CLI command line interface in detail, including the following topics:

- Accessing the CLI of the switch
- Introduction to CLI Mode
- Command Syntax Introduction
- Command Line Shortcut
- History Command

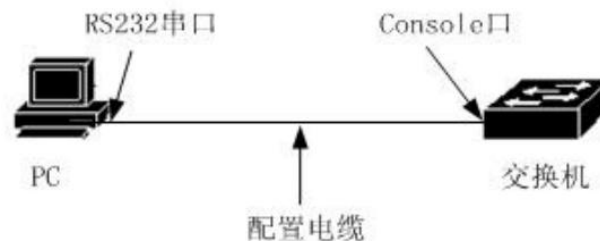
1.1 Accessing the CLI of the Switch

The CLI command line interface of the switch provides an interface for managing the switch. You can access the CLI through the Console port or through a Telnet terminal. Both methods are described below.

1.1.1 Accessing the CLI Through the Console Port

The steps are as follows:

Step 1: Connect the serial port of the PC to the Console port of the switch with the configuration cable, as shown in the following figure:



Step 2: Start a terminal emulation program on the PC (such as HyperTerminal on Windows) and configure its communication parameters. Configure the terminal parameters as follows:

Baud rate: 38400 (or 115200) (Note: subject to the actual product)

Data bits: 8

Parity: None

Stop bit: 1

Data flow control: None

Configure the HyperTerminal communication parameters as shown in the figure below:



Step 3: Start the switch. After the switch starts, the CLI prompt (Switch> by default) appears on the terminal. You can enter commands at this prompt, which gives you access to the CLI of the switch.

1.1.2 Accessing the CLI Through Telnet

You can access the switch through its ports.

The IP address of the switch port is 192.168.0.1 by default (note: the actual product shall prevail). To access the switch through a port, follow these steps:

Step 1: Connect the Ethernet port of the PC to a port of the switch with an Ethernet cable, as shown in the following figure:



Step 2: Set the IP address of the Ethernet port of the PC. The address must be in the 192.168.0.0/24 segment (for example, 192.168.0.100). Ping 192.168.0.1 to determine the connectivity between the PC and the switch.

Step 3: If the PC is connected to the switch, Telnet 192.168.0.1 enters the Telnet terminal interface, as shown in the following figure:



Step 4: If the system has no password set, the Telnet interface enters the CLI directly and the CLI prompt appears (Switch> by default). If the system has a password set, you must enter the password on the Telnet interface to enter the CLI.

Note the following two points:

- The IP address of a switch port is established on the VLAN Layer 3 interface. Before accessing the switch, you must set the IP address of a VLAN interface. The default IP address of VLAN1, 192.168.0.1 (Note: The actual product shall prevail), can be used directly. You can configure the IP address of the VLAN interface through the Console port.
- When you access the switch through a port, you can connect the PC directly to the port with an Ethernet cable, or connect through a network, as long as the PC and a VLAN of the switch can communicate with each other.

1.2 Introduction to CLI Mode

1.2.1 What CLI Mode Does

CLI mode has the following two main functions:

- It classifies users and prevents unauthorized users from using the CLI. Users are divided into two levels, that is, two categories: normal users and privileged users.
 - a. Ordinary users can view only part of the operating status of the switch and can use only the display command.
 - b. In addition to viewing the operating status of the switch, privileged users can maintain and configure the switch and change its behavior.
- It simplifies configuration of the switch.

The switch has many configurations. Putting them all in one mode would make the switch very inconvenient to use. For this reason, the CLI provides multiple modes and places similar commands in one mode, which makes them easier to understand and use. For example, VLAN-related commands are placed in VLAN configuration mode, and interface-related commands in interface configuration mode.

1.2.2 Identifying the CLI Mode

The CLI prompt identifies the CLI mode. When you use the CLI, you can tell the current CLI mode by looking at the prompt.

The CLI prompt consists of two parts: one identifies the host, and the other identifies the mode.

The host part of the CLI prompt uses the host name of the system. The host name is configurable and defaults to Switch. The CLI prompt therefore starts with Switch by default, and the CLI prompts shown later generally use the default host name.

The mode part of the CLI prompt is not configurable. Each mode has its own mode string. Some mode strings are fixed and some are variable. For example, the mode string of VLAN configuration mode is fixed, and the mode string of interface configuration mode is variable.

For example:

The CLI prompt Switch# identifies privileged mode. Switch identifies the host, and # identifies the mode.

The CLI prompt Switch(config-ge1/1)# identifies interface configuration mode, with the ge1/1 port being configured. Switch identifies the host, and (config-ge1/1)# identifies the mode.

The CLI prompt Switch(config-vlan2)# identifies interface configuration mode, with the vlan2 interface being configured. Switch identifies the host, and (config-vlan2)# identifies the mode.

1.2.3 Classification of CLI Modes

CLI modes fall into four categories: normal mode, privileged mode, global configuration mode, and configuration submenu. The configuration submenu category is made up of many CLI modes.

Normal users have access only to normal mode. Privileged users have access to all CLI modes.

Console and Telnet terminals first enter normal mode. Enter the enable command in normal mode, and you enter privileged mode after the password is verified successfully. On a Telnet terminal, a normal user can only stay in normal mode and cannot enter privileged mode. Enter configure terminal in privileged mode to change the CLI mode to global configuration mode. You enter each configuration submenu by entering the relevant command in global configuration mode.

The following table lists the main CLI modes of the switch:

Mode	Description	Prompt	Command to Enter Mode	Command to Exit Mode
Normal Mode	Provides the display command to view status information for the switch.	Switch>	The mode the terminal enters first.	There is no exit mode command on the Console terminal. On a Telnet terminal, use the exit or quit command to exit the Telnet terminal.
Privileged Mode	In addition to display commands for viewing the status information of the switch, provides commands such as debugging, version upgrade, and configuration maintenance.	Switch#	Enter the enable command in normal mode.	Use the disable command to return to normal mode. On the Console terminal, use the exit or quit command to exit to normal mode. On a Telnet terminal, use the exit or quit command to exit the Telnet terminal.
Global Configuration Mode	Provides generic commands that cannot be implemented within a configuration submode, such as the command to configure a static route.	Switch(config)#	Enter the configure terminal command in privileged mode.	Use the exit, quit, or end command to exit to privileged mode.
Interface Configuration Mode	Provides commands for configuring ports and VLAN interfaces.	Port: Switch(config-ge1/1)# VLAN interface: Switch(config-vlan1)#	Enter the interface <if-name> command in global configuration mode.	Use the exit or quit command to exit to global configuration mode, or the end command to exit to privileged mode.
VLAN Configuration Mode	Provides commands for configuring VLANs, such as commands to create and delete VLANs.	Switch(config-vlan)#	Enter the vlan database command in global configuration mode.	Use the exit or quit command to exit to global configuration mode, or the end command to exit to privileged mode.

Mode	Description	Prompt	Command to Enter Mode	Command to Exit Mode
MSTP Configuration Mode	Provides commands for configuring MSTP, such as commands to create and delete an MSTP instance.	Switch(config-mst)#	Enter the spanning-tree mst configuration command in global configuration mode.	Use the exit or quit command to exit to global configuration mode, or the end command to exit to privileged mode.
Terminal Configuration Mode	Provides commands for configuring Console and Telnet terminals, such as a timeout for the terminal.	Switch(config-line)#	Enter the line vty command in global configuration mode.	Use the exit or quit command to exit to global configuration mode, or the end command to exit to privileged mode.

1.3 Command Syntax Introduction

1.3.1 Command Composition

A CLI command consists of two parts: keywords and parameters. The first word must be a keyword, and the words that follow can be keywords or parameters. Keywords and parameters can appear alternately. A command must have a keyword, but it need not have parameters. For example, the command write has one keyword and no parameters; the command show version has two keywords and no parameters; the command VLAN <vlan-id> has one keyword and one parameter; the command instance <instance-id> VLAN <vlan-id> has two keywords and two parameters, with the keywords and parameters alternating.

1.3.2 Parameter Type

CLI commands have two types of parameters: required and optional. You must enter a required parameter when you enter the command, while an optional parameter may or may not be entered. The parameter in the command VLAN <vlan-id> is required, so you must enter it when you enter the command. The parameter in the command show interface [if-name] is optional, so it may or may not be entered.

1.3.3 Command Syntax Rule

1) The following rules must be met when describing commands with text:

Keywords are expressed directly in words.

For example, the command show version.

2) Parameters must be enclosed in < >.

For example, the command VLAN <vlan-id>

3) An optional parameter must be enclosed in [].

For example, the command show VLAN [<vlan-id>]

In this case, the < > around the parameter can be omitted, giving:

Command show VLAN [VLAN-ID]

Command show VLAN [VLAN-ID]

A required parameter cannot have [].

4) If one of several keywords or parameters must be selected, enclose the keywords or parameters with { }, separate them with |, and include a space before and after |.

For example, a command in which one of several keywords is required:

```
spanning-tree mst link-type {point-to-point | shared}
```

You must choose between point-to-point and shared.

Commands with multiple required parameters:

```
no arp {<ip-address> | <ip-prefix>}
```

Commands in which both keywords and parameters are mandatory:

Show spanning-tree mst {none|instance <0-15>}ng

5) If more than one keyword or parameter can be selected, use [] to enclose the keywords or parameters, separate them with |, and include a space before and after |.

The command is as follows:

```
debug ip tcp [recv | send]
```

The keywords recv and send may or may not be selected.

```
show ip route [<ip-address> | <ip-prefix>]  
show interface [<if-name> | switchport]
```

6) If a keyword, a parameter, or a group of keywords or parameters can be selected and entered repeatedly, add the symbol "*" after the keyword or parameter (group).

For example, ping:

ping <ip-address> [-n <count> | -l <size> | -r <count> | -s <count> | -j <count> <ip-address> / -k <count> <ip-address> | -w <timeout>]*

-j <count> <ip-address> --- Multiple IP addresses can be entered repeatedly. -k <count> <ip-address> --- Multiple IP addresses can be entered repeatedly. The entire option can also be entered repeatedly.

7) A parameter is represented by a descriptor of one or more words. If there are multiple words, separate each word with the symbol "-". Each word is lowercase.

Correct parameter notation: <vlan-id>, <if-name>, <router-id>, <count>, and so on.

Incorrect parameter notation: <1-255>, <A.B.C.D>, <WORD>, <IFNAME>, and so on.

1.3.4 Command Abbreviation

When you enter a command in the CLI interface, you can abbreviate the command keyword. The CLI supports prefix matching: as long as the word you enter matches a keyword prefix uniquely, the CLI parses it into the matching keyword. This is very convenient when you use the CLI, because you can type very few characters to complete a command. For example, you need type only shver for the show version command.

1.3.5 Syntax Help

The CLI command line interface provides syntax help, which supports help for commands and parameters at every level, as described below:

- 1) Enter ? directly in one of the CLI modes. The terminal lists the first keyword of every command in that mode and its description. For example, Switch(config)#?.
- 2) Enter the first part of a command, followed by a space and then ?. The terminal lists all the keywords or parameters of the next level and their descriptions. For example, Switch# show?.
- 3) Enter an incomplete keyword and then enter ? directly. The terminal lists all keywords that match the input prefix and their descriptions. For example, Switch# show ver?.
- 4) Enter the first part of a command UNK1, then enter a space and then the Tab key. The terminal lists all keywords of the next level. If the next level is a parameter, it is not listed.
- 5) UNK1 enter an incomplete keyword and then press the Tab key directly. If only one keyword matches the input prefix, it is completed directly. If more than one keyword matches the input prefix, the terminal lists all matching keywords.

1.3.6 Command Line Error Message

If a command you enter fails the syntax check, an error message appears on the terminal. Common error messages are shown in the following table.

Error Message	Error Cause
Invalid input or Unrecognized command	No matching keyword was found. A parameter was entered incorrectly. Too many keywords or parameters were entered.
Incomplete command	The command entry is incomplete; some keywords or parameters were not entered.
Ambiguous command	The keyword entry is incomplete. More than one keyword matches the input prefix.

1.4 Command Line Shortcut

1.4.1 Line Editing Shortcut

The CLI command line interface supports line editing shortcut keys, which make entering and editing CLI commands easier. When you enter or edit a command, you can use the line editing shortcut keys to speed up entry. The following table lists all the line editing shortcuts and the functions they perform:

Shortcut Key	Function
Ctrl+p or ↑	Previous command
Ctrl+n or ↓	Next command
Ctrl+u	Delete the entire line
Ctrl+a	Move the cursor to the beginning of the line
Ctrl+f or →	Move the cursor one space to the right
Ctrl+b or ←	Move the cursor one space to the left
Ctrl+d	Delete the character at the cursor
Ctrl+h	Delete the character before the cursor
Ctrl+k	Delete all characters at and after the cursor
Ctrl+w	Delete all characters before the cursor
Ctrl+e	Move the cursor to the end of the line
Ctrl+c	Interrupt the command line without executing it. If the CLI is in global configuration mode or a configuration submode, the CLI exits to privileged mode. If the CLI is in normal mode or privileged mode, the CLI mode stays the same, but the CLI starts a new line.
Ctrl+z	Same as Ctrl+c.
Tab	Use this key after entering an incomplete keyword. If one keyword matches the prefix entered, the keyword is completed. If more than one keyword matches the prefix entered, all matching keywords are listed. If no keyword matches, the key has no effect.

Note: The ↑, ↓, →, and ← keys are not available on some Console terminals.

1.4.2 Show Command Shortcut

Commands that begin with the show keyword are display commands. Some display commands show too much content to fit on one screen. The terminal provides a split-screen display: after displaying one screen, it waits for your input to determine the subsequent processing. The following table lists the display command shortcut keys and their functions.

Shortcut Key	Function
Space	Show the next screen
Enter	Show the next line
Ctrl+c	Interrupt command execution and exit to CLI mode
Other keys	Same as Ctrl+c.

1.5 History Command

The CLI command line interface supports a command history function. It remembers the last 20 history commands you used and saves the last commands you typed. Use `show history` to display the commands that have been entered. You can also use Ctrl+p, Ctrl+n, or the ↑ and ↓ keys to select history commands. The history command function makes entering commands easier.

Chapter 2 System Management Configuration

Before you learn how to configure the features of the switch, you need to master the basic configuration of system management and maintenance. This chapter describes that basic configuration, including the following topics:

- System Security Configuration
- System maintenance and commissioning
- System Monitoring
- Configuration File Management
- Software version upgrade

2.1 System Security Configuration

To prevent unauthorized users from invading the switch, the system provides several security measures for system management, mainly the following:

- Multi-user management control
- Enable Password Control
- TELNET Service Control
- SNMP Service Control
- HTTP Service Control

2.1.1 Multi-User Management Control

Multi-user management ensures the security of the switch system and also lets several users manage and maintain the switch at the same time. It ensures system security by giving each user a user name, a password, and a permission. When you access the switch, you must first pass user name and password verification. Verification succeeds only when the user name and password are correct and consistent. After you pass authentication, you can access the switch, but your permission limits the scope of your access.

Multi-user management divides user permissions into two levels: ordinary user and privileged user. An ordinary user can only stay in normal mode of the CLI command line interface and can use only the display command to query information about the switch. A privileged user has access to all modes of the CLI command line interface and can use all commands provided by the CLI to query information about the switch and to maintain and manage it.

The multi-user management feature applies only to Telnet terminals and does not control Console terminals. Accessing the switch through the Console terminal requires no user name and password authentication; you enter the CLI directly. Accessing the switch through a Telnet terminal requires user name and password verification, and you can access the CLI only after they are verified.

The default user name and password of the switch are both admin. The admin user must be an administrator, that is, a privileged user. It cannot be configured as an ordinary user, and the admin user cannot be deleted.

The commands related to multi-user management are shown in the following table:

Command	Description	CLI Mode
<code>username <user-name> password <key> {normal privilege}</code>	Add a user. If the specified user already exists, modify that user's password and permission. The first parameter is the user name, the second parameter is the password, and the optional parameter represents the permission: normal for a normal user, privilege for a privileged user.	Global Configuration Mode
<code>no username [user-name]</code>	Delete the user with the specified user name.	Global Configuration Mode
<code>show running-config</code>	View the current configuration of the system to see the multi-user management configuration.	Privileged Mode

2.1.2 Enable Password Control

The enable password controls the switch from normal mode to privileged mode. Before the enable password is verified, you can only view information about the switch. After it is verified, you can configure and maintain the switch.

The enable password does not depend on the user. Any user who logs in through the Console terminal or a Telnet terminal must pass enable password verification to enter privileged mode. If verification fails, the user can only stay in normal mode.

Enter the enable command in normal mode, and the terminal prompts you for the password. Enter the enable password. If it is verified successfully, the terminal enters privileged mode; otherwise, the terminal stays in normal mode. Ordinary users cannot enter privileged mode, whether or not the password is verified successfully.

The enable password is empty by default. In this case, after you enter the enable command in normal mode, the terminal enters privileged mode directly without prompting for the password.

The commands related to the enable password are shown in the following table:

Command	Description	CLI Mode
<code>enable password <key></code>	Set the enable password for the system.	Global Configuration Mode
<code>no enable password</code>	Clear the enable password of the system. The enable password is blank.	Global Configuration Mode
<code>show running-config</code>	View the current configuration of the system to see the enable password configuration.	Privileged Mode
<code>enable</code>	Interactive command to verify the enable password of the system. After successful verification, the terminal enters privileged mode.	Normal Mode

Note: For system security, the administrator needs to set the enable password of the system.

2.1.3 TELNET Service Control

In some cases, the administrator does not need to manage the switch remotely and only needs to manage it locally through the Console terminal. To improve system security and prevent unauthorized users from logging in remotely through a Telnet terminal, the administrator can turn off the Telnet service. The Telnet service is on by default.

The commands related to Telnet service control are shown in the following table:

Command	Description	CLI Mode
<code>security-manage telnet enable</code>	Turn on the Telnet service.	Global Configuration Mode
<code>security-manage telnet disable</code>	Turn off the Telnet service.	Global Configuration Mode
<code>security-manage telnet number <1-100></code>	The number parameter can range from 1 to 100. The default is 5.	Global Configuration Mode
<code>security-manage telnet access-group <1-99></code> (Note: subject to the actual product)	Specify an ACL group to turn on source IP address control. If the specified ACL group does not exist or is not a standard ACL group, the source IP address is not controlled.	Global Configuration Mode
<code>no security-manage telnet access-group</code>	Turn off source IP address control.	Global Configuration Mode
<code>show security-manage</code>	View the configuration of the service control.	Privileged Mode

2.1.4 SNMP Service Control

SNMP service control turns the SNMP service on or off and also controls, through ACLs, the IP addresses that access the switch.

The commands related to SNMP service control are shown in the following table:

Command	Description	CLI Mode
<code>security-manage snmp enable</code>	Turn on the SNMP service.	Global Configuration Mode
<code>security-manage snmp disable</code>	Turn off the SNMP service.	Global Configuration Mode
<code>Security-manage snmp access-group <1-99></code> (Note: subject to the actual product)	Specify an ACL group to turn on source IP address control. If the specified ACL group does not exist or is not a standard ACL group, the source IP address is not controlled.	Global Configuration Mode
<code>no security-manage snmp access-group</code>	Turn off source IP address control.	Global Configuration Mode
<code>show security-manage</code>	View the configuration of the service control.	Privileged Mode

2.1.5 HTTP Service Control

HTTP service control turns the HTTP service on or off and also controls, through ACLs, the IP addresses that access the switch.

The commands related to HTTP service control are shown in the following table:

Command	Description	CLI Mode
<code>security-manage http enable</code>	Turn on the HTTP service.	Global Configuration Mode
<code>security-manage http disable</code>	Turn off the HTTP service.	Global Configuration Mode
<code>security-manage http access-group <1-99> (Note: subject to the actual product)</code>	Specify an ACL group to turn on source IP address control. If the specified ACL group does not exist or is not a standard ACL group, the source IP address is not controlled.	Global Configuration Mode
<code>no security-manage http access-group</code>	Turn off source IP address control.	Global Configuration Mode
<code>show security-manage</code>	View the configuration of the service control.	Privileged Mode

2.1.6 HTTPS Service Control

HTTPS service control turns the HTTPS service on or off and also controls, through ACLs, the IP addresses that access the switch.

The commands related to HTTPS service control are shown in the following table:

Command	Description	CLI Mode
<code>security-manage https enable</code>	Turn on the HTTP service.	Global Configuration Mode
<code>security-manage https disable</code>	Turn off the HTTP service.	Global Configuration Mode
<code>security-manage https access-group <1-99> (Note: subject to the actual product)</code>	Specify an ACL group and turn on source IP address control. If the specified ACL group does not exist or is not a standard ACL group, the source IP address is not controlled.	Global Configuration Mode
<code>no security-manage https access-group</code>	Turn off source IP address control.	Global Configuration Mode
<code>show security-manage</code>	View the configuration of the service control.	Privileged Mode

2.1.7 SSH Service Control

Traditional network service programs, such as FTP, pop, and Telnet, are inherently insecure, because they transmit passwords and data over the network in plaintext. People with ulterior motives can easily intercept these passwords and data. The security verification methods of these service programs also have a weakness:

they are vulnerable to "man-in-the-middle" attacks. In a "man-in-the-middle" attack, the "middleman" pretends to be the real server to receive the data you send to the server, and then pretends to be you to send the data to the real server. Serious problems arise once the "middleman" tampers with the data transmitted between you and the server. SSH encrypts all transmitted data, so a "man-in-the-middle" attack cannot be carried out, and it also prevents DNS spoofing and IP spoofing. Another advantage of SSH is that the transmitted data is compressed, which speeds up transmission. SSH has many functions. It can replace Telnet, and it can also provide a secure "channel" for FTP, pop, and even PPP.

The commands related to SSH are shown in the following table:

Command	Description	CLI Mode
<code>security-manage ssh enable</code>	Turn on the SSH service.	Global Configuration Mode
<code>security-manage ssh disable</code>	Turn off the SSH service.	Global Configuration Mode
<code>security-manage ssh access-group <1-99></code>	Specify an ACL group and turn on source IP address control. If the specified ACL group does not exist or is not a standard ACL group, the source IP address is not controlled.	Global Configuration Mode
<code>no security-manage ssh access-group</code>	Turn off SSH source IP address control.	Global Configuration Mode
<code>ip sshd auth-retries <times></code>	Set the maximum number of verification attempts the client can make	Global Configuration Mode
<code>ip sshd silence-period <seconds></code>	Set the waiting time after the client exceeds the maximum number of verifications	Global Configuration Mode
<code>show ip sshd</code>	View the configuration information of sshd	Privileged Mode
<code>show security-manage</code>	View the configuration of the service control.	Privileged Mode

2.2 System Maintenance and Commissioning

The basic system maintenance and commissioning functions mainly include the following:

- Configure the system's clock
- Configure Terminal Timeout Properties
- System Reset
- View system information
- Network connectivity debugging
- Traceroute Debug

2.2.1 Configure the System's Clock

The switch provides a real-time clock function. You can set or view the current clock with a command. The system clock is powered internally so that the real-time clock keeps running when the system is powered off, and the clock does not need to be reset after the system starts.

The clock is set when the switch leaves the factory, so you do not need to set it again. If you find that the time is not accurate, you can reset the clock.

The commands related to the system clock are shown in the following table:

Command	Description	CLI Mode
<code>set date-time <year> <month> <day> <hour> <minute> <second></code>	To set the current clock of the system, enter the year, month, day, hour, minute, and second parameters.	Privileged Mode
<code>show date-time</code>	Display the current clock of the system.	Normal Mode, Privileged Mode

2.2.2 Configure Terminal Timeout Properties

For terminal security, a terminal exits after a certain time with no key input. Console terminals and Telnet terminals are handled differently. On a Console terminal, when the terminal times out, the CLI mode exits to normal mode. On a Telnet terminal, when the terminal times out, the Telnet connection is interrupted and the Telnet terminal exits. The default terminal timeout is 10 minutes. You can also set the terminal to never time out.

The commands related to terminal timeout are shown in the following table:

Command	Description	CLI Mode
<code>exec-timeout <minutes> [seconds]</code>	Set the timeout time of the terminal. If the parameters are all 0, the terminal never times out.	Terminal Configuration Mode
<code>no exec-timeout</code>	Set the terminal timeout back to the default, which is 10 minutes.	Terminal Configuration Mode
<code>show running-config</code>	View the current configuration of the system to see the terminal timeout configuration.	Privileged Mode

2.2.3 System Reset

The system provides one reset method:

- Reset the switch

The commands related to system reset are shown in the following table:

Command	Description	CLI Mode
reset	Reset the switch.	Privileged Mode

2.2.4 View System Information

The system provides a variety of display commands for viewing the operating status and information of the system. Only a few display commands commonly used for system maintenance are listed here, as shown in the following table:

Command	Description	CLI Mode
show version	Displays the version number of the system and the time at which the file was compiled and linked.	Normal Mode, Privileged Mode
show snmp system information	Displays basic information about the system, including how long the system has been running since it started.	Normal Mode, Privileged Mode
show history	Displays a list of the most recently entered commands on the CLI command line.	Normal Mode, Privileged Mode

2.2.5 Network Connectivity Debugging

To debug the connectivity between the switch and another device in the network, run the ping command on the switch and ping the IP address of the other device. If the switch receives a ping response from the other device, the two ends are connected. Otherwise, the two ends cannot communicate.

The switch implements the ping command and also supports many options on it, which you can use for more precise and complex debugging.

The ping command is shown in the following table:

Command	Description	CLI Mode
ping <ip-address> [-n <count> -l <size> -w <timeout>]*	It can be used without any options, or with one or more options. Without any options, it is the simplest ping command. Press Ctrl+c while the command is running to interrupt it.	Privileged Mode

2.2.6 Traceroute Debug

To debug which intermediate devices the switch passes through when it communicates with another device on the network, run the trace-route command on the switch. Specify the IP address of the other device, and all the paths in between are displayed while the command runs.

The switch implements the trace-route command and also supports a number of options on it, which you can use for more precise and complex debugging.

The trace-route command is shown in the following table:

Command	Description	CLI Mode
<code>trace-route <ip-address> [-h <maximum-hops> -j <count> <ip-address>* -w <timeout>]*</code>	It can be used without any options, or with one or more options. Without any options, it is the simplest trace-route command. Press Ctrl+c while the command is running to interrupt it.	Privileged Mode

2.2.7 Telnet Client

The switch provides a Telnet client function, through which you can remotely access other devices.

Command	Description	CLI Mode
<code>telnet <ip-address></code>	The parameter is the IP address of the target device	Privileged Mode

2.2.8 SSH Client

The series switches provide an SSH client function, through which you can remotely access other devices.

Command	Description	CLI Mode
<code>ssh <ip-address> [user-name]</code>	Access other devices remotely with the SSH client	Privileged Mode

2.3 Configuration File Management

The configuration is divided into the current configuration and the initial configuration. The current configuration is the configuration in use while the system is running, and it is stored in the memory of the system. The initial configuration is the configuration used when the system starts, and it is stored in the FLASH of the system, that is, in the configuration file. When you run a configuration command, the current configuration of the system is modified. The current configuration is written into the initial configuration for the next startup only after the save command is executed. After the system starts, if you make no configuration, the current configuration information is the same as the initial configuration information.

The current configuration and the initial configuration use the same format, the format of command line text, which is intuitive and easy to read. The format of the configuration file has the following characteristics:

- The configuration file is a text file.
- Everything saved is a command.
- Only the non-default configuration is saved; the default configuration is not saved.
- Commands are organized according to the CLI mode. Commands in the same CLI mode are grouped together to form a segment, and the segments are separated by "!". In global configuration mode, commands with the same or similar functions are grouped into a segment and separated by "!".
- Commands in a configuration submode are preceded by a space, while commands in global configuration mode do not need a leading space.

- End the configuration with "end".

Configuration file management mainly includes the following:

- View configuration information
- Save Configuration
- Delete the configuration file
- Configuration file upload and download

2.3.1 View Configuration Information

Viewing configuration information includes viewing the current and the initial configuration of the system. The initial configuration is the configuration file in FLASH. When the configuration file does not exist in FLASH, the system starts with the default configuration. In that case, if you view the initial configuration of the system, the system reports that the configuration file does not exist.

The commands for viewing configuration information are shown in the following table:

Command	Description	CLI Mode
<code>show running-config</code>	View the current configuration of the system.	Privileged Mode
<code>show startup-config</code>	View the initial configuration of the system.	Privileged Mode

2.3.2 Save Configuration

When you modify the current configuration of the system, you need to save it to the configuration file so that the configuration still exists after the next startup; otherwise, the configuration information is lost after a restart. Saving a configuration means saving the current configuration to the initial configuration.

The commands to save the configuration are shown in the following table:

Command	Description	CLI Mode
<code>write</code>	Save the current configuration.	Privileged Mode

Note: Use this command to save the configuration after you configure the switch; otherwise, the configuration is lost after the system restarts.

2.3.3 Delete the Configuration File

When you want the initial configuration of the system to return to the default configuration, you can delete the configuration file. Deleting the configuration file has no effect on the current configuration. If you want the current configuration of the system to return to the default configuration, you must restart the switch. Be careful when deleting configuration files; otherwise, the configuration is lost.

The command to delete the configuration file is shown in the following table:

Command	Description	CLI Mode
<code>delete startup-config</code>	Delete the system's configuration file.	Privileged Mode

2.3.4 Configuration File Upload and Download

To keep the configuration file safe, you can use a command to upload it to a PC for backup. If the system configuration is lost or modified abnormally and you want to return to the original configuration, you can download the original configuration file from the PC to the switch. The downloaded configuration file has no effect on the current configuration of the system; the configuration takes effect after the switch restarts. Configuration files can also be uploaded and downloaded through the WEB. For the specific operations, refer to the WEB Operation Manual. The commands for downloading the configuration file are as follows:

Command	Description	CLI Mode
<code>upload configure <ip-address> <file-name></code>	The configuration file is uploaded to the PC. The first parameter is the IP address of the PC, and the second parameter is the file name under which the configuration file is stored on the PC.	Privileged Mode
<code>download configure <ip-address> <file-name></code>	The configuration file is uploaded to the PC. The first parameter is the IP address of the PC, and the second parameter is the file name under which the configuration file is stored on the PC.	Privileged Mode

The configuration file is downloaded with the TFTP protocol: the TFTP client software runs on the switch, and the TFTP server software runs on the PC. To download the configuration file, proceed as follows:

Step 1: Build a network environment.

Step 2: Start the TFTP server software on the PC and set the directory where the configuration file is stored.

Step 3: Save the configuration on the switch.

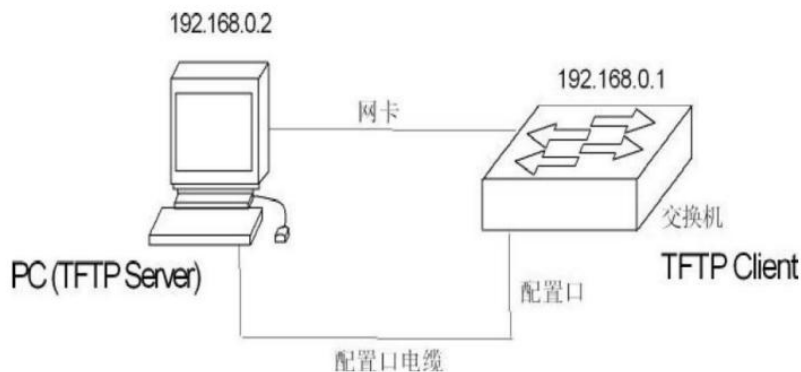
Step 4: Execute the configuration file upload command on the switch to back up the configuration file to the PC.

Step 5: When the switch needs the configuration file on the PC, execute the configuration file download command on the switch to download the configuration file from the PC to the switch.

Step 6: The switch must be restarted for the configuration to take effect.

Example: a switch that has been configured with VLANs and interface addresses needs to upload and download configuration files.

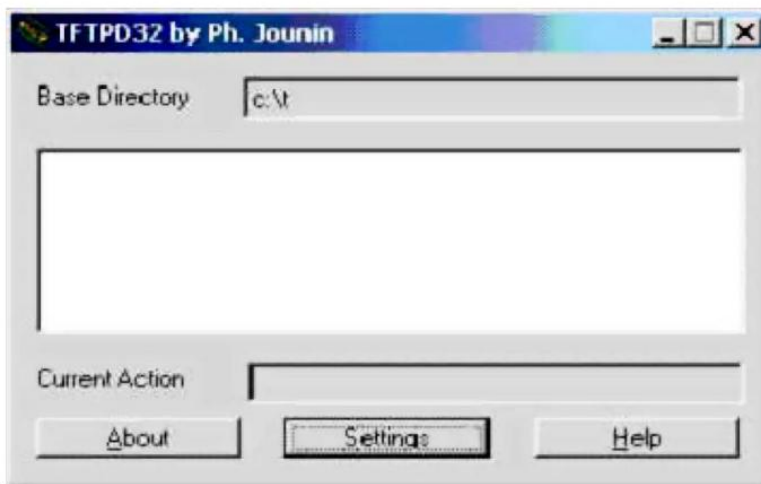
Step 1: Set up the network environment as shown below.



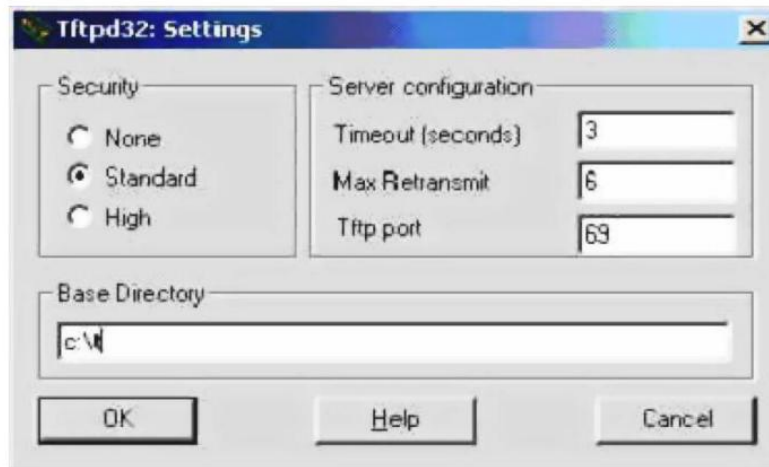
The configuration port of the switch is connected to a configuration terminal through a cable, and to a PC through a network cable. Install TFTP Server on the PC and configure the IP address of the Ethernet port of the PC. Here it is assumed that the IP address of the PC is 192.168.0.2. Next, configure the IP address of the switch, assuming that it is 192.168.0.1. Ensure connectivity between the PC and the switch.

Step 2: Start TFTP Server and configure the TFTP Server parameters.

Run TFTP Server. The window interface is shown below:



Then set the directory where the configuration files are backed up. Click the [Settings] button to enter the setting interface, as shown in the following figure:



Enter the file path in Base Directory. Click the [OK] button to confirm. Step 3: Execute the write command on the switch to save the current configuration to the configuration file.

Step 4: Back up the file to the PC by executing the command Switch# upload configuration 192.168.0.2 beifen.CFG.

Step 5: If necessary, download the backup file to the switch by executing the command Switch# download configuration 192.168.0.2 beifen.CFG.

Step 6: For the downloaded configuration file to take effect, you must restart the switch by executing the command Switch# reset.

2.4 Software Version Upgrade

The switch supports online upgrade of the software version. The upgrade is done with the TFTP tool.

2.4.1 Commands for Software Version Upgrade

1) Upgrade the switch file of the switch in global configuration mode with the following command:

download switch <ip-address> <file-name>

Here, <ip-address> is the IP address of the PC running the TFTP server, and <file-name> is the name of the switch file saved on the TFTP server.

2) Upgrade the kernel file of the switch in global configuration mode with the following command:

download kernel <ip-address> <file-name>

Here, <ip-address> is the IP address of the PC running the TFTP server, and <file-name> is the name of the kernel file saved on the TFTP server.

3) Upgrade the patch file of the switch in global configuration mode with the following command:

download patch <ip-address> <file-name>

Here, <ip-address> is the IP address of the PC running the TFTP server, and <file-name> is the name of the patch file saved on the TFTP server.

4) Upgrade the uboot file of the switch in global configuration mode with the following command:

download uboot <ip-address> <file-name>

Here, <ip-address> is the IP address of the PC running the TFTP server, and <file-name> is the name of the uboot file saved on the TFTP server.

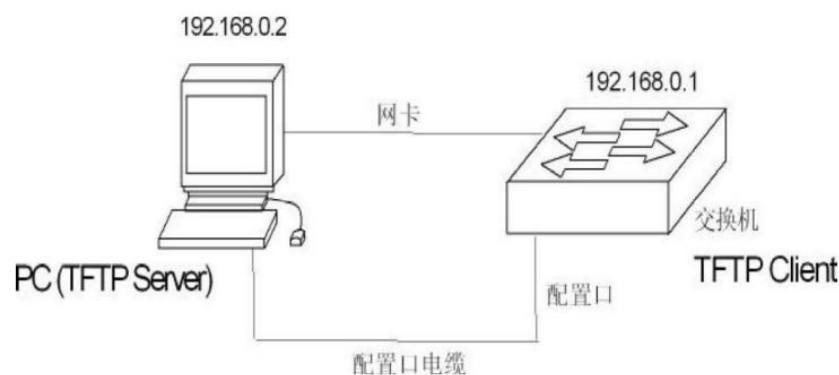
Do not power off the switch during the upgrade; otherwise, the switch files may be damaged and the switch may fail to start. After the download is complete, restart the switch to run the newly downloaded file program. The entire upgrade process takes a few minutes. Please be patient.

The software version can also be upgraded through the WEB. For the specific operations, refer to the WEB Operation Manual.

2.4.2 Software Upgrade Process

The steps to upgrade the switch file and other files are as follows:

Step 1: Set up the upgrade environment, as shown in the figure below.

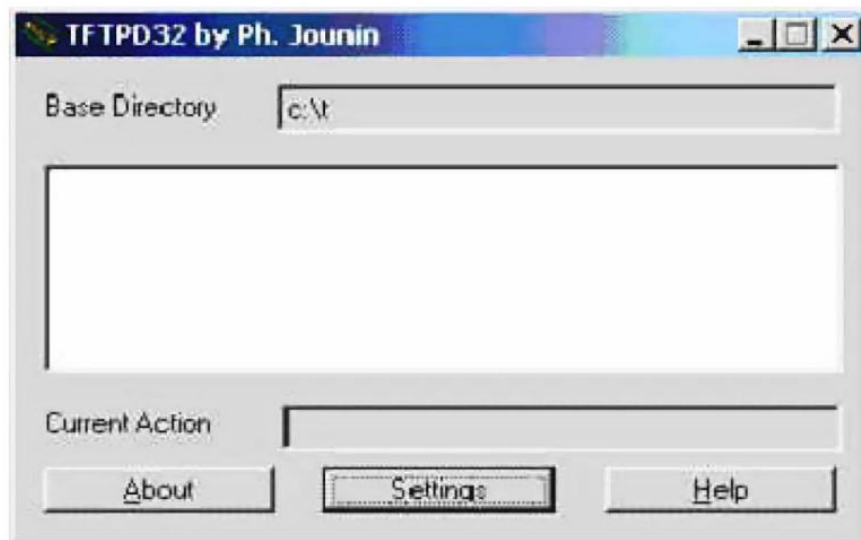


The setup process is as follows:

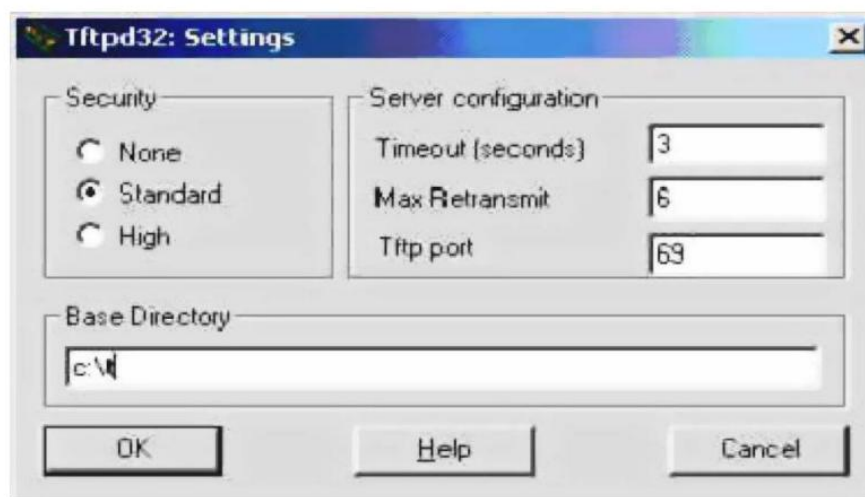
- Connect the Console port of the switch to a configuration terminal (PC) with a cable.
- Install TFTP Server on the PC.
- Copy the new switch file and other files to a path on the PC, assuming that the path is C:\t.
- Configure the IP address of the Ethernet port of the PC. It is assumed that the IP address of the PC is 192.168.0.2.
- Configure the IP address of the switch, assuming that the switch IP address is 192.168.0.1.

Step 2: Run TFTP Server and configure the TFTP server.

First, run TFTP Server. The TFTP32 window interface is shown in the figure below:



Then set the TFTP Server file directory. After starting TFTP Server, reset the TFTP Server file directory and copy the switch file and other files to be loaded into this directory. To do this, click the [Settings] button to display the TFTP32 setting interface, as shown below.



Enter the file path in Base Directory. Click the [OK] button to confirm.

Step 3: Upgrade the file.

First, connect a port on the switch to the PC running the TFTP Server program with an Ethernet cable. Then use the ping command to check connectivity between the host and the switch.

Then enter the command at the HyperTerminal Switch# prompt:

```
Switch# download switch 192.168.0.2 switch.img, press Enter, and wait for the completion of file upgrade.
```

Software is updating. Please wait and don't powerdown!

.....

Updating is completed. Do you wish to reset?[Y/N]

After the file is transferred, the system asks whether you want to restart the switch. In general, select 'Y' to restart the switch, because the system upgrade takes effect only after the switch restarts. If your configuration file is not saved, you can select 'N' first and then restart the switch after you finish saving and any other operations.

```
Switch#
```

Note:

The switch must not be powered down during the upgrade process.

Step 4: Reboot the switch.

```
Switch# reset
```

Chapter 3 Configure Port

This chapter describes the configuration related to ports, including the following topics:

- General configuration of the port
- Configure MIRROR
- Configure STORM-CONTROL
- Configure flow-control
- Configure port bandwidth
- Configure TRUNK

3.1 General Configuration of the Port

The administrator can control the users under a port by configuring the port of the switch. If the users under a port are not allowed to access the network, the administrator can close the port. This section describes the general configuration of ports, mainly:

- Opening and closing of port
- Rate configuration of the port
- Displays information about the port

3.1.1 Opening and Closing of Port

The port of the switch is open by default. If the administrator wants to prevent the users under a port from accessing the network, the administrator can close the port.

The following command turns on the administrative state of the port in interface configuration mode:

```
no shutdown
```

For example, open the administrative state of port ge1/1:

```
Switch(config-ge1/1)#no shutdown
```

The following command turns off the administrative state of the port in interface configuration mode:

Shutdown

For example, close the administrative state of port ge1/1:

```
Switch(config-ge1/1)#shut
```

3.1.2 Rate Configuration of the Port

The default rate configuration for all ports is autonegotiate.

The following command configures the rate of the port in interface configuration mode:

```
speed {autonegotiate |full-1000 |full-100 |full-10 |half-100 |half-10 }
```

autonegotiate --- Auto-negotiation full-1000 --- Full-duplex gigabit full-100 --- Full-duplex 100M full-10 --- Full-duplex 10M half-100 --- Half-duplex 100M half-10 --- Half-duplex 10M

For example, configure the rate of port 1/1 as full-duplex 100M:

```
Switch(config-ge1/1)# speed full-10
```

3.1.3 Displays Information About the Port

The following command displays information for one or more ports in normal or privileged mode:

```
show interface [if-name]
```

For example, show the information for port ge1/1:

```
Switch# show interface ge1/1
```

For example, display information for all ports:

```
Switch# show interface
```

3.2 Configure MIRROR

Port mirroring is a useful feature for monitoring the traffic received and sent by one or more ports. It uses a mirroring port to monitor the packets received and sent by one or more ports. The switch supports port mirroring, and the mirroring port can listen to the incoming and outgoing data of other ports. One mirror port can listen on multiple ports at the same time. This section describes the configuration of MIRROR, including the following:

- Configure the listening and listened ports of MIRROR
- Display MIRROR configuration

3.2.1 Configure the Listening and Listened Ports of MIRROR

To configure the monitoring port, the administrator enters interface configuration mode and sets the monitored port. For example, to set port ge1/1 to monitor port ge1/2, the administrator enters port ge1/1 and types the command:

```
Switch(config-ge1/1)# mirror interface ge1/2 direction both
```

Port ge1/1 is then set as the listening port, and ge1/2 is set as the listened port.

The command to set the listening port is as follows:

```
Switch(config-ge1/1)#mirror interface <if-name> direction {both | receive | transmit}
```

In this case, port ge1/1 is set as the monitoring port and <if-name> is set as the monitored port. The following { both | receive | transmit } indicates the monitoring direction: receive monitors the received packets; transmit monitors the transmitted packets; both listens to all packets sent and received. For example:

```
Switch(config-ge1/1)#mirror interface ge1/2 direction both
```

This indicates that port ge1/1 is set to listen to the packets sent and received by port ge1/2.

To set up multiple listening ports, execute multiple commands.

In interface configuration mode, the administrator can cancel the monitored port with the following command:

```
Switch(config-ge1/1)#no mirror interface <if-name>
```

Here, <if-name> is the port that is no longer listened to. For example:

```
Switch(config-ge1/1)# no mirror interface ge1/2
```

This indicates that port ge1/1 is set to no longer listen to the packets on port ge1/2. When all listened ports are canceled, the listening ports are cleared.

3.2.2 Display MIRROR Configuration

The administrator can view the MIRROR configuration that has been set by using the following command in normal mode or privileged mode:

```
Switch# show mirror
```

Note the following points:

- A port cannot be set as both a listening port and a listened port.
- There can be only one listening port, but there can be more than one listened port.

3.3 Configure STORM-CONTROL

In real networks, a NIC card can send unicast, multicast, and broadcast packets at high speed, which can cause network failures. In this case, the suppression function on the switch is especially important, because it prevents packets from flooding the network and causing congestion. All ports of the switch support the suppression function for broadcast packets, multicast packets, and

DLF packets.

This section describes the configuration of STORM-CONTROL in detail, including the following topics:

- Default configuration
- Broadcast Suppression Configuration
- Multicast suppression configuration
- DLF Suppression Configuration
- Display the STORM-CONTROL configuration

3.3.1 Default Configuration

The switch supports setting broadcast, multicast, and dlf switches for each port, and the three settings share the same rate limit. Broadcast packet suppression is on by default on the port, at a rate of 64K, to prevent a broadcast storm from forming on the network. DLF packets and multicast packets are not suppressed by default.

3.3.2 Broadcast Suppression Configuration

The following command configures broadcast suppression for this port in interface configuration mode:

```
storm-control broadcast
```

The following command cancels broadcast suppression for this port in interface configuration mode:

```
no storm-control broadcast
```

3.3.3 Multicast Suppression Configuration

The following command configures multicast suppression for this port in interface configuration mode:

```
storm-control multicast
```

The following command cancels multicast suppression for this port in interface configuration mode:

```
no storm-control multicast
```

3.3.4 DLF Suppression Configuration

The following command configures DLF suppression for this port in interface configuration mode:

```
storm-control dlf
```

The following command removes DLF throttling for this port in interface configuration mode:

```
no storm-control dlf
```

3.3.5 Suppression Rate Configuration

The following command configures the throttle rate for this port in interface configuration mode:

```
storm-control ratelimit <1-1024000>
```

3.3.6 Display the STORM-CONTROL Configuration

The following command displays the STORM-CONTROL configuration in normal or privileged mode:

```
show storm-control
```

3.4 Configure STORM-CONSTRAIN

The port traffic threshold control function controls packet storms on the Ethernet. A port with this function enabled periodically detects the unicast, multicast, and broadcast packet traffic arriving at the port. If the traffic of a given packet type exceeds the preset upper threshold, you can decide through configuration whether to block or shut down the port, and whether to record log information.

When the traffic of a given packet type exceeds the preset upper threshold for that type, the system provides two processing methods:

- (1) Block mode: If the unicast, multicast, or broadcast traffic on the port is greater than the upper threshold, the port stops forwarding those packets and enters the blocking state, but it still counts the traffic of this packet type. When the traffic of this packet type falls below the preset lower threshold, the port resumes forwarding.
- (2) Shutdown mode: If the unicast, multicast, or broadcast traffic on the port is greater than the upper threshold, the port is shut down and the system stops forwarding all packets. You can restore the port state by executing the no shutdown command or by unconfiguring the port traffic threshold.

Note: A given packet type can be suppressed either by this function or by the storm suppression function of the Ethernet port, but the two functions cannot be configured at the same time; otherwise the suppression effect is uncertain. For example, you cannot configure the unicast packet traffic threshold control function and the unicast storm suppression function of the port at the same time.

The CLI configuration commands are as follows:

Command	Description	CLI Mode
<code>storm-constrain (broadcast multicast unicast)</code>	Storm control for broadcast, multicast or unknown unicast packets under the interface	Interface Configuration Mode
<code>min-rate <1-1488100> max-rate <1-1488100></code>		
<code>no storm-constrain (broadcast multicast unicast all)</code>	Cancel storm control	Interface Configuration Mode
<code>storm-constrain action (block shutdown)</code>	Configure the storm control action. By default, no storm control is performed on packets	Interface Configuration Mode
<code>no storm-constrain action</code>	Cancel the configured storm control action	Interface Configuration Mode
<code>storm-constrain enable (trap)</code>	Turn on logging during storm control	Interface Configuration Mode
<code>no storm-constrain enable (log)</code>	Turn off logging during storm control	Interface Configuration Mode
<code>storm-constrain interval <6-180></code>	Configure the storm control detection interval. By default, the storm control detection interval is 5 seconds	Interface Configuration Mode
<code>no storm-constrain interval</code>	Restore the storm control detection interval to the default value	Interface Configuration Mode
<code>no storm-constrain</code>	Remove the storm control feature from the interface	Interface Configuration Mode
<code>show storm-constrain</code>	View storm control information for all interfaces	Privileged Mode
<code>show storm-constrain interface IFNAME</code>	View the storm control information for an interface	Privileged Mode

Configuration description:

(1) Description table for the interface storm control information display

Item	Description
Interface	Interface name
type	Packet type: (1) broadcast-broadcast packet; (2) multicast-multicast packet; (3) unicast-unicast packet
rate	Min-low threshold; Max-high threshold
action	Storm control actions, including (1) block-block the packet; (2) shutdown-shut down the interface
punish-status	Packet status of the current interface, including (1) block-when the rate is greater than max-rate and the storm control action is to block packets, the status is blocking packets; (2) Normal-normal forwarding; (3) shutdown-when the rate is greater than max-rate and the storm control action is to shut down the interface, the status is interface shut down
log	Log switch status, on/off
interval	The storm control detection interval, in seconds. The default value is 5 seconds
last-punish-time	Time of the last storm control penalty

(2) You can control storm packets and prevent flooding by executing the storm-constrain action command to configure the storm control action, and the storm-constrain command to configure the high and low thresholds of storm control. During the storm control detection interval, when the average rate of broadcast, multicast, or unicast packets received on the interface is greater than the specified high threshold, storm control blocks the packets or shuts down the interface according to the configured action. When the storm control action is to block packets, the interface returns to normal forwarding if the traffic falls below the minimum threshold. When the storm control action is to shut down the interface, the interface cannot recover automatically; you must execute the no shutdown command manually, or recover it by canceling the port storm control shutdown action configuration.

(3) When the port traffic exceeds the upper threshold, or falls from the upper threshold to the lower threshold, log/trap information is output.

3.5 Configure Flow-Control

FLOW-CONTROL prevents packet loss when a port is blocked. In half-duplex mode, flow control is implemented by backpressure technology, so that the information source reduces its transmission rate. In full-duplex mode, flow control follows the IEEE 802.3 X standard, and the blocked port sends a "Pause" packet to the information source to stop it from sending.

This section describes the configuration of FLOW-CONTROL in detail, mainly including the following contents:

- Default configuration
- Set port flow control
- Turn off port flow control

- Display flow control information

3.5.1 Default Configuration

The switch supports setting flow control for sending and receiving on each port. By default, flow control is not turned on for a port.

3.5.2 Set Port Flow Control

The following command turns on port flow control in interface configuration mode:

```
flowcontrol
```

3.5.3 Turn Off Port Flow Control

The following command turns off flow control on the transmit and receive sides of the port in interface configuration mode:

```
no flowcontrol
```

3.5.4 Display Flow Control Information

The following command displays flow control information for all ports in normal or privileged mode:

```
show flowcontrol
```

The following command displays flow control information for a port in normal or privileged mode:

```
show flowcontrol interface <if-name>
```

Where <if-name> is the name of the port whose flow control information you want to query.

3.6 Configure Port Bandwidth

Port bandwidth control controls the rate at which a port sends and receives.

This section describes the configuration of port bandwidth in detail, mainly including the following contents:

- Default configuration
- Set port send or receive bandwidth control
- Remove port send or receive bandwidth control
- Show bandwidth control for port configuration

3.6.1 Default Configuration

The switch supports setting the transmit and receive bandwidth for each port separately. By default, a port has no bandwidth control.

3.6.2 Set Port Send or Receive Bandwidth Control

The following command sets the port transmit or receive bandwidth control in interface configuration mode:

```
portrate {egress | ingress} <rate>
```

egress indicates bandwidth control of the transmitted packets.

ingress indicates bandwidth control of the received packets.

<rate> indicates the value of the bandwidth to be set, in the range of 1-1024000, in kbits.

3.6.3 Remove Port Send or Receive Bandwidth Control

The following command removes bandwidth control from a port in interface configuration mode:

```
no portrate {egress | ingress}
```

egress indicates that the bandwidth control of the transmitted packets is canceled.

ingress indicates that the bandwidth control of the received packets is canceled.

3.6.4 Show Bandwidth Control for Port Configuration

The following command displays the bandwidth control of the port configuration in normal or privileged mode:

```
show portrate interface <if-name>
```

Where <if-name> is the name of the port whose bandwidth control information you want to query.

3.7 Configure TRUNK

TRUNK aggregates multiple ports into one logical port, which can be used to increase bandwidth, provide redundant backup connections, and balance load. When the trunk group is used as the output logical port, the switch selects a port from the port group to send the packet according to the aggregation policy you set. The port and aggregation policy of the trunk group are configured by software, but the data stream is forwarded by hardware.

All ports in a trunk group must be configured at the same speed and in full-duplex mode. The switch supports up to 32 groups of TRUNKs, with up to 8 TRUNK members per group. Note that each port can belong to only one TRUNK group.

The LACP protocol is based on the IEEE 802.3ad standard. LACP exchanges information with the peer through the LACPDU (Link Aggregation Control Protocol Data Unit).

When an interface in the aggregation group enables the LACP protocol, the interface informs the peer of its system LACP protocol priority, system MAC, port LACP protocol priority, port number, and operation Key by sending LACPDUs. After receiving the LACPDUs, the peer compares the information in the LACPDUs with the information received on other interfaces to select the interface that can be in the Selected state, so that both sides agree on the interface in the Selected state.

The action Key is a combination of configurations that aggregation control generates automatically at link aggregation time, based on certain configurations of the member ports: port rate, duplex mode, up/down status, VLANs allowed on the port, port default VLAN ID, and link type of the port (that is, Trunk, Hybrid, or Access type). In an aggregation group, member ports in the Selected state have the same action Key.

This section describes the configuration of TRUNK in detail, mainly including the following contents:

- LACP protocol configuration
- Configuration of the trunk group
- TRUNK member port configuration
- TRUNK load balancing policy configuration
- Display of TRUNK

3.7.1 LACP Protocol Configuration

Command	Description	CLI Mode
<code>lacp system-priority <1-65535></code>	Set the lacp system priority	Global Configuration Mode
<code>no lacp system-priority</code>	Restore the system priority to the default value 32768	Global Configuration Mode
<code>lacp max-active-link-number <1-8></code>	Set the lacp activation aggregation port upper limit	Global Configuration Mode
<code>no lacp max-active-link-number</code>	Restore the lacp activation aggregation port upper limit to the default value 8	Global Configuration Mode
<code>lacp port-priority <1-65535></code>	Set the lacp port priority	Interface Configuration Mode
<code>no lacp port-priority</code>	Restore the port priority to the default value 32768	Interface Configuration Mode
<code>lacp timeout (short long)</code>	Set the lacp port timeout. The default is long timeout	Interface Configuration Mode
<code>show lacp summary</code>	Briefly display the status of all lacp aggregates	Privileged Mode
<code>show lacp detail</code>	Display the status of all lacp aggregates	Privileged Mode
<code>show lacp <1-8></code>	Display the details of the lacp aggregation port	Privileged Mode
<code>show lacp port IFNAME</code>	Display the details of the lacp port	Privileged Mode
<code>show lacp system-id</code>	Display the lacp system status	Privileged Mode

Command	Description	CLI Mode
<code>show lacp counter <1-8></code>	Display the statistics of the lacp aggregation port	Privileged Mode
<code>show lacp counter</code>	Display the statistics of all lacp aggregation ports	Privileged Mode
<code>clear lacp <1-8> counters</code>	Clear the statistics of the lacp aggregation port	Privileged Mode
<code>clear lacp counters</code>	Clear the statistics of all lacp aggregation ports	Privileged Mode

3.7.2 Configuration of the Trunk Group

The following command creates a manual TRUNK group in global configuration mode:

```
trunk <trunk-id>
```

This creates a TRUNK group. The <trunk-id> value ranges from 1 to 32 and indicates the ID number of the TRUNK group to create. Up to 8 groups of TRUNKs can be configured. After the group is created, the interface name of the TRUNK is trunk + ID number. For example, the interface name of a TRUNK group with a group ID of 1 is trunk1. In configuration mode, use the "interface trunk + ID number" command to enter interface configuration mode and then operate on the TRUNK group. For example, use the command interface trunk1 to enter the interface mode of TRUNK 1 and configure TRUNK 1.

The following command creates a static LACP trunk group in global configuration mode:

```
trunk <1-32> dynamic
```

The following command removes a trunk group in global configuration mode:

```
no trunk <trunk-id>
```

Before deleting a trunk group, make sure the trunk group has no member ports.

3.7.3 TRUNK Member Port Configuration

The following command adds a new TRUNK group member port in interface configuration mode:

```
trunk interface IFNAME (passive|)
```

<if-name> is the name of the port to add to the TRUNK group, which must be a layer 2 interface. Each group of TRUNKs can add up to 8 layer 2 interfaces. If the trunk group is a static LACP trunk group, the added interface defaults to the active state and can also be configured to the passive state.

The following command removes all member ports of the trunk group in interface configuration mode:

```
no trunk interface
```

The following command removes the specified TRUNK group member port in interface configuration mode:

```
no trunk interface <if-name>
```

You can use this command multiple times to delete multiple ports that are members of the TRUNK group.

3.7.4 TRUNK Load Balancing Policy Configuration

The following command sets the load balancing policy for TRUNK in interface configuration mode:

```
trunk load-balance {dst-mac | dst-ip | src-dst-mac | src-dst-ip | src-mac |
```

src-ip} dst-mac-----Balancing strategy based on destination MAC dst-ip-----Balancing strategy based on destination IP src-dst-mac---Balancing strategy based on source MAC and destination

MAC src-dst-ip-----Balancing strategy based on source IP and destination IP src-mac-----Balancing strategy based on source MAC src-ip-----Balancing strategy based on source IP

The following command sets the default TRUNK load balancing policy in interface configuration mode:

```
no trunk load-balance
```

The default port load balancing policy is src-dst-mac (a balancing policy based on source and destination MAC).

3.7.5 Display of TRUNK

The following command displays all TRUNKs in normal or privileged mode: show trunk

The following command displays the configuration of a specified TRUNK in normal or privileged mode:

```
show trunk <trunk-id>
```

Where <trunk-id> is the ID number of the trunk group to query.

3.8 Configure Jumbo Frame

3.8.1 Introduction to Jumbo Frame

To enable a port to receive jumbo frames, set the port to support a specific jumbo frame length.

3.8.2 Jumbo Frame Configuration

Configure the port to support the jumbo frame length. In configuration mode, enter port configuration mode, such as interface ge1/1, and execute the following command:

```
Switch(config-ge1/1)#jumbo frame 2000
```

Display the jumbo frame length supported by the port

```
Switch#show jumbo frame ge1/1
```

Port Jumbo frame(bytes)

ge1/1 2000

3.9 Configuring Redundant Port

In some special cases, such as when you need to ensure the stability of servers linked to the network, the redundant port feature of the switch provides two ports linked to the server and ensures that the server has only one LINK UP port linked to the network at a time. When LINK DOWN occurs on one port, the system enables the other port.

When a port in a redundant port group is in the LINK UP state, it is in the Active state. If a port in a redundant port group is in the LINK DOWN state, it is in the Disable state.

This section describes the configuration of redundant ports, including the following:

- Configuration of redundant port
- Display of redundant port

3.9.1 Configuration of Redundant Port

The switch can be configured with 8 groups of redundant ports. One group of redundant ports can be configured with only 2 ports. One port can be configured in only one redundant port group.

A redundant port group can be configured with one primary-port and one secondary-port. When you enable a redundant port group:

- (1) When both ports are in the LINK UP state, the primary-port is set to the Active state and the secondary-port is set to the Disable state;
- (2) If only one port is in the LINK UP state, that port is set to the Active state and the other port is in the Disable state;
- (3) Otherwise, both ports are in the Disable state.

If a LINK DOWN event occurs on a port that is in the Active state, the other port attempts to enter the Active state.

Another configuration parameter is force-switch, which applies when the secondary-port is in the Active state and the primary-port is in the Disable state. If a LINK UP event occurs on the primary-port, force-switch decides whether to switch back to the state where the primary-port is Active and the secondary-port is Disable. If force-switch is configured to enable, the switchover is forced; otherwise the port state of the original redundant port group is retained.

Command	Description	CLI Mode
redundant-port <1-8> primary-port IFNAME secondary-port IFNAME [force-switch]	Configure a set of redundant ports. <1-8> is the group number; primary-port IFNAME is the interface name of the primary port; secondary-port IFNAME is the interface name of the secondary port; force-switch specifies whether the forced switchover is enabled.	Global Configuration Mode
redundant-port <1-8> force-switch	Enable the forced switchover for the redundant port.	Global Configuration Mode
no redundant-port <1-8>	Delete a redundant port group.	Global Configuration Mode
no redundant-port <1-8> force-switch	Turn off the forced switchover for the redundant port.	Global Configuration Mode

3.9.2 Display of Redundant Port

Command to display the redundant port

Command	Description	CLI Mode
show redundant-port	Display the configuration of all redundant port groups in the system	Privileged Mode

3.10 DLDP Configuration

DLDP (device link detection protocol) is a layer 2 protocol used to monitor the physical configuration of an Ethernet link connected by optical fiber or twisted pair. A unidirectional link is one where transmission works in only one direction; for example, this device can send data to the peer and the peer receives it, but this device cannot receive the data the peer sends. DLDP detects this situation, shuts down the corresponding interface, and sends a warning message. Unidirectional links can cause many problems, especially with spanning tree, where they can cause loops. Note: DLDP requires support from both devices on the link to operate properly.

DLDP supports two modes of operation: normal mode (the default) and aggressive mode.

Normal mode: In this mode, DLDP detects unidirectional links and marks the port as undetermined to generate system logs. In other words, normal mode shuts down a port only if it can explicitly determine that the associated link has been faulty for an extended period of time.

Aggressive mode: In this mode, DLDAP detects unidirectional links. It tries to rebuild the link and sends DLDAP messages for 8 seconds. If there is no DLDAP echo reply during this period, the port is placed in the errdisable state.

Command	Description	CLI Mode
<code>dldap enable</code>	Enable the DLDAP function globally	Global Configuration Mode
<code>dldap message time <time></code>	DLDAP message sending interval	Global Configuration Mode
<code>dldap port</code>	Enable DLDAP on the port	Interface Configuration Mode
<code>dldap aggressive</code>	Enable aggressive mode on the port. The default is normal mode	Interface Configuration Mode
<code>show dldap <ifname></code>	View port DLDAP information	Privileged Mode

3.11 Configure LLDP

There are more and more kinds of network devices, and their configurations are complex. A standard information exchange platform is needed so that devices from different manufacturers can find each other and exchange their system and configuration information on the network.

LLDP (Link Layer Discovery Protocol) was created for this purpose and provides a standard way of link layer discovery. Information such as the main capability, management address, device identifier, and interface identifier of the local device can be organized into different TLVs (Type/Length/Value) and encapsulated in an LLDPDU (Link Layer Discovery Protocol Data Unit) sent to the neighbors directly connected to it. After receiving this information, the neighbors save it in a standard MIB (Management Information Base) so that the network management system can query it and judge the communication status of the link.

This section describes the configuration of LLDP, mainly including the following:

- Configuration of LLDP
- Display of LLDP

3.11.1 Configuration of LLDP

The LLDP port has four modes of operation:

TxRx: send and receive LLDP messages.

Tx: send LLDP messages but do not receive them.

Rx: receive LLDP messages but do not send them.

Disable: neither send nor receive LLDP messages.

When the LLDP working mode of the port changes, the port initializes the protocol state machine. To avoid frequent changes of the working mode causing the port to initialize continuously, you can configure a delay time for port initialization, so that when the working mode of the port changes, the initialization is performed after a period of time.

Command	Description	CLI Mode
<code>lldp global enable</code>	LLDP global enable command.	Global Configuration Mode
<code>lldp hold-multiplier <num></code>	LLDP TTL multiplier.	Global Configuration Mode
<code>lldp timer [<reinit-delay><time>][<tx-delay><time>][<tx-interval><time>]</code>	Configure the various timers for LLDP.	Global Configuration Mode
<code>lldp enable</code>	Enable LLDP on the interface	Interface Configuration Mode
<code>lldp admin-status{disable rx tx rxtx}</code>	Configure the operating mode of the LLDP port.	Interface Configuration Mode
<code>lldp check-change-interval <time></code>	Configure the time interval for refreshing interface information.	Interface Configuration Mode
<code>lldp management-address <A.B.C.D></code>	Configure the interface LLDP management address	Interface Configuration Mode
<code>lldp tlv-enable{dot1-tlv dot3-tlv med-tlv}</code>	Configure the interface LLDP extended capability set switch	Interface Configuration Mode

3.11.2 Display of LLDP

Commands for LLDP

Command	Description	CLI Mode
<code>show lldp configuration [ifname]</code>	Display lldp configuration information	Privileged Mode
<code>show lldp local-information [ifname]</code>	Display lldp local information	Privileged Mode
<code>show lldp neighbor-information [ifname]</code>	Display lldp neighbor information	Privileged Mode
<code>show lldp statistics [ifname]</code>	Display lldp message statistics	Privileged Mode
<code>show lldp status [ifname]</code>	Display lldp status information	Privileged Mode

Chapter 4 Configuring Port-Based Security

This chapter describes port-based MAC security configuration, mainly including the following contents:

- Introduction
- MAC binding configuration
- MAC filtering configuration
- Port learning limit configuration
- Protection port configuration

4.1 Introduction

Port-based MAC security provides four functions: MAC binding, MAC filtering, port learning control, and port protection. They improve the security of layer 2 forwarding.

MAC binding binds a MAC address and a port together, so a specified MAC address can access the network only on a specified port. At the same time, the port allows only these bound MAC addresses to access the network. A port can bind multiple MAC addresses at the same time. MAC binding can be applied to a given port at the same time as 802.1X. This feature is useful for devices that do not support 802.1X or are not suited to using it, such as printers and file servers.

MAC filtering prevents specified MAC addresses from accessing the network. Its main purpose is to keep unauthorized devices off the network. When a MAC address is configured for MAC filtering, that MAC address cannot access the network on any port of the switch, nor can it receive packets destined for it. As with MAC binding, a port can be configured with multiple MAC addresses for MAC filtering at the same time. In practice, if virus software attacks the network through forged MAC addresses, you can use MAC filtering in addition to ACLs to control the attack of forged packets.

Port learning control controls the number of MAC addresses that a port can learn dynamically. If a port specifies the number of MAC addresses it can learn dynamically, then when the number of MAC addresses learned by the port equals the number configured for the port, no new MAC addresses are learned and packets for these new MAC addresses are discarded.

Note that the MAC address referred to here is actually MAC + VID; this is not repeated in the rest of this chapter. In addition, the MAC binding function and 802.1X can be configured on one port at the same time, and MAC filtering and the port learning limit can be configured on one port at the same time, but the MAC binding function and 802.1X cannot be configured on the same port as MAC filtering and the port learning limit.

4.2 MAC Binding Configuration

MAC binding configuration supports manual and automatic MAC address binding. Manual binding means you enter the MAC addresses and ports one by one through commands. Automatic binding reads the existing entries of the port in the layer 2 hardware forwarding table and binds the MAC addresses directly. The command to read the layer 2 hardware table is `show bridge fdb`.

Configuration Command

Command	Description	CLI Mode
switchport-security mac-bind HHHH.HHHH.HHHH vlan <1-4094> qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7}	Manually bind a MAC address to an interface and configure the priority queue of the entry.	Interface Configuration Mode
switchport-security mac-bind auto-conversion number <1-8191> qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7}	Automatically convert a specified number of MAC addresses of an interface into a MAC binding configuration and configure the priority queue for the entry.	Interface Configuration Mode
switchport-security mac-bind auto-conversion vlan <1-4094> qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7}	Automatically convert the MAC addresses of a specified VLAN on an interface into a MAC binding configuration and configure the priority queue for the entry.	Interface Configuration Mode
show port-security mac-bind [IFNAME]	Display the MAC binding configuration	Privileged Mode

Notice:

A MAC address binding can be invalid or fail for the following reasons:

The port is already configured with 802.1x;

The port is already configured with MAC filtering or a port learning limit;

The MAC address is already bound to another port, or MAC filtering is configured;

The L2 table of the switch is full.

4.3 MAC Filtering Configuration

MAC filtering configuration supports manual and automatic MAC address binding. Configuring MAC filtering manually means you enter the MAC address to filter and the port to bind one by one through commands. Configuring MAC filtering automatically reads the existing entries of the port in the layer 2 hardware forwarding table and configures MAC filtering directly. The command to read the layer 2 hardware table is show bridge fdb.

Configuration Command

Command	Description	CLI Mode
<code>switch port-security mac-filter HHHH.HHHH.HHHH vlan <1-4094></code>	Manually configure MAC filtering for an interface	Interface Configuration Mode
<code>switch port-security mac-filter auto-conversion number <1-8191></code>	Automatically convert a specified number of MAC addresses of an interface into a MAC filtering configuration	Interface Configuration Mode
<code>switch port-security mac-filter HHHH.HHHH.HHHH vlan <1-4094></code>	Automatically convert the MAC addresses of a specified VLAN on an interface into a MAC filtering configuration	Interface Configuration Mode
<code>show port-security mac-filter [IFNAME]</code>	Display the MAC binding configuration	Privileged Mode

Notice:

A MAC filtering configuration can be invalid or fail for the following reasons: The port is already configured with MAC binding or has the 802.1x protocol function enabled;

The MAC address is already bound to another port, or MAC binding is configured;

The L2 table of the switch is full.

4.4 Port Learning Limit Configuration

The switch can configure the maximum number of dynamically learned addresses per port. If a port is configured with a number of MAC addresses to learn dynamically, the port can learn only that number of MAC addresses. When the number of MAC addresses exceeds this number, the port cannot learn or forward them.

Without a configured learning limit, a port can learn up to 8191 MAC addresses.

Configuration Command

Command	Description	CLI Mode
<code>switchport port-security learn-limit <0-8191></code>	Configure the number of MAC addresses that an interface can learn.	Interface Configuration Mode
<code>no switchport port-security learn-limit</code>	Delete the number of MAC addresses an interface can learn.	Interface Configuration Mode
<code>show port-security learn-limit [IFNAME]</code>	Show the port learning configuration	Privileged Mode

Configuration examples

Configure port ge1/5 to learn only 7 MAC addresses

```
Switch#configure terminal
Switch(config)interface ge1/5
Switch(config-ge1/5)switchport port-security learn-limit 7
```

Notice:

Port configuration can be invalid or fail for the following reasons:

The port has MAC binding configured or the 802.1x protocol function enabled.

4.5 Protection Port Configuration

4.5.1 Introduction to Protection Port

To achieve layer 2 isolation between packets, you can divide ports into different VLANs, but this wastes VLAN resources. Configuring ports as protection ports isolates ports in the same VLAN and provides a safer and more flexible networking scheme.

Each port can be configured as a protected port. Protected ports cannot communicate with each other and can communicate only with unprotected ports. There are two application modes:

1. Only one port is not configured as a protection port, and all other ports are isolated;
2. To prevent insecure ports from sniffing the data of other ports (connected to the server), set these ports as protected ports.

4.5.2 Protection Port Configuration

Configure the port as a protection port. In configuration mode, enter port configuration mode, such as interface ge1/1, and execute the following command:

```
Switch(config-ge1/1)#switchport port-security protect
```

Display the ports that are configured as protection ports

```
Switch#show port-security protect
```

Port Port protected

-----ge1/1 ON

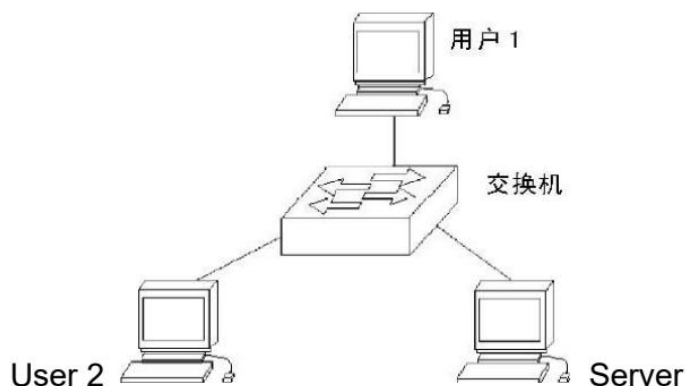
Example of a typical protection port configuration

Networking requirements

User 1, user 2, and the server are connected to switch ports ge1/1, ge1/2, and ge1/3, respectively

User 1 and user 2 belong to the same VLAN. User 1 and user 2 cannot communicate with each other, but they can communicate with the server.

Network Diagram



Configuration step

```
Switch>enable
Switch#config terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport port-security protect
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport port-security protect
Switch(config-ge1/2)#exit
Switch(config)#exit
Switch#show port-security protect
```

Port Port protected

-----ge1/1 ON ge1/2 ON

Chapter 5 Configure Port IP and MAC Bindings

This chapter describes port IP and MAC binding configuration, mainly including the following contents

- Introduction
- IP and MAC binding configuration
- Configuration examples
- Configuration troubleshooting

5.1 Introduction

Configuring IP and MAC binding on layer 2 switch ports is a static defense against ARP attacks. ARP attackers attack users by sending ARP packets with false MAC addresses, which causes the user's local ARP cache table to be overwritten with the attacker's MAC address, so normal data flows to the attacker. You can filter ARP attack packets effectively by statically binding the user IP address and MAC address in the switch port configuration command.

Besides preventing ARP spoofing, the IP MAC binding function also guarantees a one-to-one mapping between IP and MAC: one IP can correspond to only one MAC, and one MAC can correspond to only one IP. If the access device modifies this mapping, it cannot communicate on this network. The 802.1X anti-ARP spoofing function and the DHCP snooping protocol are dynamic implementations of this function.

The four functions IP MAC binding, ACL, 802.1X anti-ARP spoofing, and DHCP SNOOPING all use the same system resource, CFP. Pay attention to whether the CFP resource is exhausted during configuration. The compatibility relationship between them was defined at design time. See the following table:

	IP MAC Binding	ACL	802.1x	DHCP SNOOPING
IP MAC Binding	Compatible	Incompatible	Compatible	Compatible
ACL	Incompatible	Compatible	Incompatible	Incompatible
802.1x	Compatible	Incompatible	Compatible	Incompatible
DHCP SNOOPING	Compatible	Incompatible	Incompatible	Compatible

CFP is a limited hardware resource, and each port can be configured with only 16 IP MAC binding entries on average. Therefore, in a network with many access hosts, if only a few ports or a few IP and MAC addresses need to be controlled, use the static IP MAC binding function. This avoids the data forwarding failure caused by exhaustion of the CFP resource.

In addition, whether to use the 802.1X or the DHCP SNOOPING protocol depends on the situation. If static IP address configuration is used and the 802.1X protocol is used to access the network, 802.1X anti-ARP spoofing must be used for it to be effective. If dynamic IP address acquisition is used, use the DHCP SNOOPING protocol.

5.2 IP and MAC Binding Configuration

IP and MAC binding is configured in interface mode

Configure port IP and MAC bindings

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#ip mac-bind A.B.C.D MAC
```

Remove port IP and MAC binding

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#no ip mac-bind A.B.C.D MAC
```

Display the configuration

Display the binding table entries for all ports

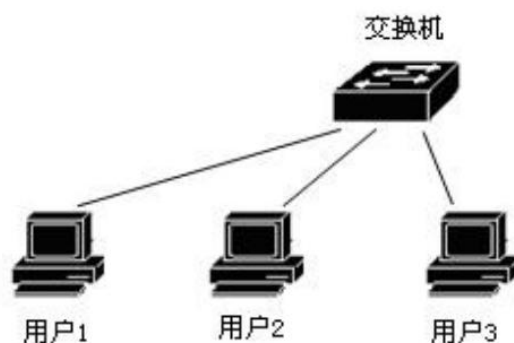
```
show ip mac-bind
```

Display the binding table entries for an interface

```
show ip mac-bind IFNAME
```

5.3 Configuration Examples

There are user 1, user 2, and user 3 in the network, and the IP and MAC of each user are bound on the port to defend against ARP attacks.



```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#ip mac-bind 192.168.1.100 0011.5b34.42ad
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#ip mac-bind 192.168.1.101 0011.6452.135d
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#ip mac-bind 192.168.1.102 0011.804d.a246
Switch(config-ge1/3)#end
Switch#show ip mac-bind
```

[ge1/1] sum: 1

MAC IP

0011.5b34.42ad 192.168.1.100

[ge1/2] sum: 1

MAC IP

0011.6452.135d 192.168.1.101

[ge1/3] sum: 1

MAC IP

0011.804d.a246 192.168.1.102

```
Switch#show ip mac-bind ge1/1
```

[ge1/1] sum: 1

MAC IP

0011.5b34.42ad 192.168.1.100

```
Switch#show running-config
```

!

```
spanning-tree mst configuration
```

!

Interface vlan1

```
ip address 10.10.10.1/24
```

!

```
interface ge1/1
ip mac-bind 192.168.1.100 0011.5b34.42ad
```

!

```
interface ge1/2
ip mac-bind 192.168.1.101 0011.6452.135d
```

!

```
interface ge1/3
ip mac-bind 192.168.1.102 0011.804d.a246
```

!

```
line vty
```

!

```
end
```

5.4 Configuration Troubleshooting

If the IP MAC binding configuration fails, the cause may be one of the following:

- 1 System CFP resources are exhausted.
- 2 The current interface is configured with ACL filtering.
- 3 The configured interface is a layer 3 interface or a TRUNK interface.

Chapter 6 Port Loop Detection

This chapter includes the following contents:

- Introduction
- Protocol principle
- Configuration introduction

6.1 Introduction

When a loop appears under a port of the switch, it causes a broadcast storm under that port, and the source MAC addresses of all broadcast packets are learned on the looped port, which causes device forwarding to fail.

6.2 Protocol Principle

The Ethernet loop detection protocol (Ethernet Loopback Detection, referred to below as ELD) detects loops through the exchange of data packets and blocks the port on which the loop occurs. The ELD protocol is a port-based protocol and can detect only loops that occur on the port itself.

6.2.1 Detection Process

When a port enables the ELD protocol, a timer period starts on the port, and a loop detection packet is sent when the timer expires. If the port receives the loop detection packet it sent within a timer period, a loop is considered to exist on the port. The switch then blocks the loop on the port and clears the FDB table of the port.

If a port is a port member of multiple VLANs, the port automatically sends loop detection packets to all VLANs. This means the port automatically detects loops in all VLANs to which it belongs.

6.2.2 Recovery Mode

As described above, when a port loop occurs, the port is blocked. The ELD protocol has two recovery modes that you can configure: automatic recovery and manual recovery.

Automatic recovery means that when a port is blocked, the ELD protocol starts a recovery timer. When the timer expires, it performs the reverse operation of blocking the loop and starts the loop detection timer on the port again.

Manual recovery means that after the port is blocked, the protocol no longer starts a timer to recover the port, and you must enter the command to perform the reverse operation of blocking the loop.

6.2.3 Protocol Security

The ELD protocol is vulnerable to attack in the network: a user can send an ELD protocol packet, built to the ELD packet format, to a port on which the ELD protocol is enabled, causing the port to be blocked when there is no loop and producing a wrong decision.

The ELD protocol uses two strategies to prevent such attacks and minimize errors.

First, the ELD protocol is a non-interactive protocol. Because it does not depend on other devices, the data packet itself can be encrypted simply. Here, the ELD protocol packet is sent with a key, and a user cannot disguise the protocol packet without the key.

The second strategy mainly prevents an attacker from reflecting the protocol packet through a packet capture attack. You can configure the format of the data packet the switch receives in a given period to prevent the attack; this requires your configuration.

6.3 Configuration Introduction

The ELD protocol is implemented on a port basis, and there is no command to enable it uniformly.

6.3.1 Global Configuration

The global configuration sets the uniform properties of the protocol.

Command	Description	Mode
<code>loop-detection detection-time <1-65535></code>	Configure the time period for loop detection. The time must be less than half the recovery time period. The default value is 5 seconds.	Global Configuration Mode
<code>loop-detection resume-time <10-65535></code>	Configure the time period for automatic recovery. The automatic recovery time must be greater than 2 times the loop check time. This configuration takes effect if automatic recovery is enabled. The default recovery time is 600 seconds.	Global Configuration Mode
<code>loop-detection protocol-safety</code>	Enable protocol security checking, which is off by default.	Global Configuration Mode
<code>loop-detection respond-packets</code>	Configure the number of packets that must be received in a given time. This configuration takes effect if protocol security checking is enabled. The default is 10.	Global Configuration Mode

6.3.2 Interface Configuration

The interface configuration is the configuration for each port.

Command	Description	Mode
Loop-detection enable	Enable the ELD protocol on a port.	Interface Configuration Mode
Loop-detection resume	Recover manually and restart the loop check.	Interface Configuration Mode
loop-detection resume-mode {automation manual}	Configure the recovery mode, either manual recovery or automatic recovery. The default is automatic recovery.	Interface Configuration Mode
loopback-detection shutdown-mode {no-shutdown shutdown}	Configure whether the port shuts down when a loop occurs.	Interface Configuration Mode

6.3.3 Display Configuration

show loop-detection [ifname]

Display all configurations of the protocol and the configuration of an interface.

Chapter 7 Configuring VLAN

VLAN is an important concept in the switch and is widely used in practice. It is the basis for dividing multiple networks internally. VLAN is short for Virtual Local Area Network, a network that logically groups multiple devices together regardless of their physical location. Each VLAN is a logical network that has all the features and attributes of a traditional physical network. Each VLAN is a broadcast domain. Broadcast packets are forwarded only within a VLAN and cannot cross VLANs. Data traffic between VLANs must be forwarded through layer 3.

This chapter includes the following contents:

- Introduction to VLAN
- VLAN configuration
- Example VLAN configuration

7.1 Introduction to VLAN

This section introduces VLAN in detail, mainly including the following contents:

- Benefits of VLAN
- VLAN ID
- VLAN port membership type
- Default VLAN for the port
- VLAN mode of the port
- VLAN trunk
- Forwarding of traffic within a VLAN
- Subnet of VLAN

7.1.1 Benefits of VLAN

VLANs greatly extend the size of the physical network. A traditional physical network can have only a very small scale, accommodating at most a few thousand devices, while a physical network divided by VLANs can accommodate tens of thousands or even hundreds of thousands of devices. VLANs have the same functions and attributes as traditional physical networks.

Using VLANs has the following benefits:

- VLANs can effectively control the traffic in the network.

In a traditional network, all broadcast packets are delivered to all devices, whether they are needed or not, which increases the load on the network and the devices. VLANs organize devices into a logical network according to your needs. A VLAN is a broadcast domain. Broadcast packets are transmitted only within the VLAN and do not cross the VLAN. Dividing the network into VLANs effectively controls the traffic in the network.

- VLANs can improve network security.

A device in a VLAN can perform layer 2 communication only with devices in the same VLAN. To communicate with another VLAN, the traffic must be forwarded through layer 3. If layer 3 forwarding between VLANs is not established, communication between VLANs is not possible at all, which isolates the VLANs and keeps the data in each VLAN secure. For example, if the R&D department of a company does not want to share data with the marketing department, the R&D department can establish one VLAN and the marketing department another, and no layer 3 communication channel is established between the two VLANs.

- VLANs facilitate the mobility of devices.

When a device in a conventional network moves from one location to another and belongs to a different network, the network configuration of the moved device must be modified, which is inconvenient. A VLAN is a logical network, and devices that are not in the same physical location can be assigned to the same network. When a device moves, it can still belong to the same VLAN. In this way, the moved device does not need any configuration changes.

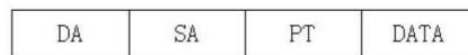
7.1.2 VLAN ID

Each VLAN has an identification number called the VLAN ID. VLAN IDs range from 0 to 4095, where 0 and 4095 are not used, so only 1 to 4094 are actually valid. The VLAN ID uniquely identifies a VLAN.

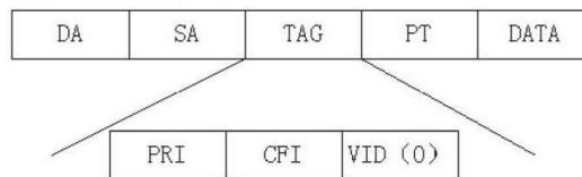
The switch supports 4094 VLANs. When creating a VLAN, select a VLAN ID in the range 2 to 4094. The switch creates VLAN 1 by default, and VLAN 1 cannot be deleted.

Three types of data frames are transmitted within a VLAN in the network: untagged data frames, tagged data frames with a VID of 0, and tagged data frames with a VID other than 0. The following figure shows the three data frame formats.

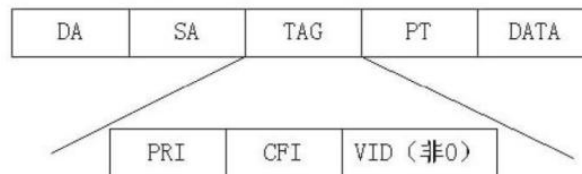
不带标记的数据帧



带标记的数据帧，但VLAN ID为0



带标记的数据帧，但VLAN ID非0



Inside the switch, all data frames are tagged. If an untagged data frame enters the switch, the switch adds a tag to the frame and selects a VLAN ID value to fill in the tagged VID. If a tagged data frame with a VID of 0

enters the switch, the switch selects a VLAN ID value to fill in the tagged VID. If a data frame tagged with a VID other than 0 enters the switch, the frame is unchanged.

7.1.3 VLAN Port Membership Type

The switch supports port-based VLANs and 802.1Q-based VLANs. A VLAN includes two types of port members: untagged and tagged. A VLAN can include both untagged and tagged port members.

A VLAN can have no port members, or one or more port members. When a port belongs to a VLAN, it can be an untagged member or a tagged member of the VLAN.

A port can be a tagged or untagged member of one or more VLANs. If a port is a tagged member of two or more VLANs, it is called a VLAN trunk port. A port can be an untagged member of one or more VLANs and a tagged member of one or more other VLANs.

7.1.4 Default VLAN for the Port

A port has one and only one default VLAN, which determines the VLAN to which an untagged packet, or a tagged packet with a VID of 0, belongs when it is received on the port. The default VLAN is also known as the port VID, or PVID. By default, the default VLAN for the port is 1.

7.1.5 VLAN Mode of the Port

Ports have three VLAN modes: ACCESS mode, TRUNK mode, and HYBRID mode. Specify the VLAN mode of the port before you configure the VLAN of the port.

A port in ACCESS mode is an access port that faces the user directly. The port can be an untagged member of only one VLAN. The default VLAN is the VLAN you specify. When a port is an untagged member of only one VLAN, you can specify ACCESS as the port's VLAN mode.

A port in TRUNK mode is a trunk port directly connected to a switch. The port can be a tagged member of one or more VLANs, but it cannot be an untagged member of any VLAN. The port's default VLAN is 1 and cannot be changed.

A port in HYBRID mode is a trunk port directly connected to a switch that can be a tagged member of one or more VLANs and an untagged member of one or more VLANs. The default VLAN for this port can be changed.

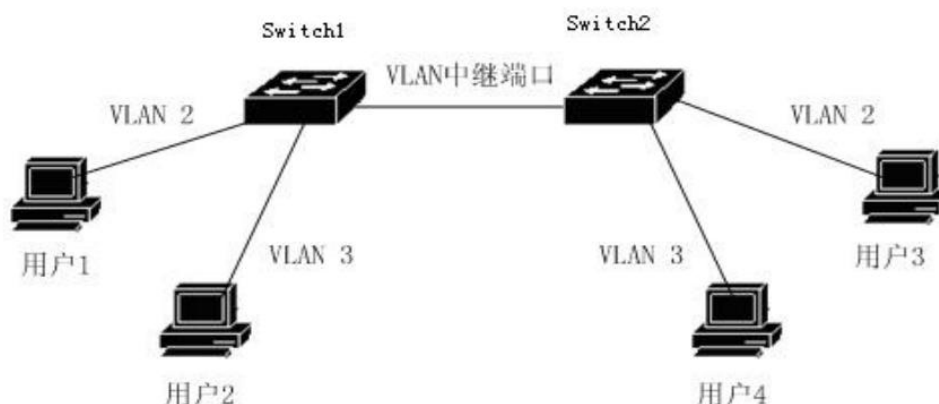
In practice, you can select the VLAN mode of the port according to the situation.

7.1.6 VLAN Trunk

A port that is a tagged member of two or more VLANs is called a VLAN trunk port. Two switches can be connected by VLAN trunk ports, so that two or more common VLANs can be divided between the two switches.

The following figure is an example of a VLAN trunk. Two switches are connected by VLAN trunk ports, which are the trunk ports of VLAN 2 and VLAN 3. Each switch is divided into two VLANs, VLAN 2 and VLAN 3, and each VLAN has a user. Thus, User 1 can communicate with User 3, and User 2 can communicate with User 4, but User 1 and User 3 cannot communicate with User 2 and User 4.

7.1.7 Forwarding of Traffic Within a VLAN



When a switch receives a packet on a port, it performs layer 2 forwarding in the following steps:

- Determine the VLAN to which the packet belongs.
- Determine whether the packet is a broadcast packet, a multicast packet, or a unicast packet.
- Determine the output ports according to the packet type (zero, one, or more output ports are allowed); if there is no output port, discard the packet.
- Determine whether the outgoing packet is tagged, according to the membership type of the output port in the VLAN.
- Send the packet out the output port.

1) How to determine the VLAN to which a packet belongs:

If a packet is received with a tag and the VID field in the tag is non-zero, the VLAN to which the packet belongs is the VID value in the tag.

If a packet is received without a tag, or with a tag whose VID value is 0, the VLAN to which the packet belongs is the port's default VLAN.

2) How to determine the type of packet:

A received packet is a broadcast packet if its destination MAC address is FF:FF:FF:FF:FF:FF.

A received packet is a multicast packet if it is not a broadcast packet and bit 40 of the destination MAC address is 1.

A packet is unicast if it is neither a broadcast packet nor a multicast packet.

3) How to determine the output port of a packet:

If the incoming packet is a broadcast packet, all member ports of the VLAN to which the packet belongs are the output ports for the packet. If the incoming packet is a multicast packet, the switch first searches the layer 2 hardware multicast forwarding table according to the destination multicast MAC address and the VLAN to which the packet belongs. If a matching multicast entry is found, the output ports of the packet are the ports common to (the AND of) the output ports in the multicast entry and the member ports of the VLAN to which

the packet belongs; if there is no common port, the packet is dropped. If no matching multicast entry is found in the layer 2 hardware multicast forwarding table, the output ports are determined according to the forwarding mode of the layer 2 hardware multicast forwarding table: if the multicast forwarding mode is unregistered multicast forwarding, the multicast packet is treated as a broadcast and all member ports of the VLAN to which the packet belongs are its output ports; if the multicast forwarding mode is registered forwarding, there is no output port and the packet is dropped.

If the incoming packet is a unicast packet, the layer 2 hardware forwarding table is searched according to the destination MAC address and the VLAN to which the packet belongs. If a matching entry is found, the output ports of the packet are the ports common to (the AND of) the output ports in the entry and the member ports of the VLAN to which the packet belongs. If there is no common port, the packet is discarded. If no matching entry is found in the layer 2 hardware forwarding table, the packet is treated as a broadcast packet, and all member ports of the VLAN to which it belongs are the output ports of the packet.

4) Send packets:

After the output ports of the incoming packet are determined, the packet is sent from all output ports. If an output port is an untagged member of the VLAN to which the packet belongs, the packet is sent out that output port untagged.

If an output port is a tagged member of the VLAN to which the packet belongs, the packet is sent out that output port with a tag whose vid value is the VLAN to which the packet belongs.

7.2 VLAN Configuration

This section describes the VLAN configuration in detail, mainly including the following contents:

- Create and delete VLAN
- Configure the VLAN mode of the port
- VLAN configuration in ACCESS mode
- VLAN configuration in TRUNK mode
- VLAN configuration in HYBRID mode
- VLAN subnet configuration
- Viewing VLAN information

7.2.1 Create and Delete VLAN

Before creating and deleting VLANs, use the `vlan database` command in global configuration mode to enter VLAN configuration mode, where VLANs are created and deleted.

The system creates VLAN 1 by default, and you cannot delete VLAN 1. The commands for creating and deleting VLANs are shown in the following table:

Command	Description	CLI Mode
<code>vlan <vlan-id></code>	Create a VLAN. If the VLAN already exists, nothing is done; otherwise the VLAN is created. The parameter ranges from 2 to 4094.	VLAN Configuration Mode
<code>no vlan <vlan-id></code>	Delete a VLAN. If the VLAN does not exist, nothing is done; otherwise the VLAN is deleted. The parameter ranges from 2 to 4094.	VLAN Configuration Mode

7.2.2 Configure the VLAN Mode of the Port

Before configuring the VLAN of the port, specify the VLAN mode of the port, which is ACCESS mode by default. The commands to specify the VLAN mode of a port are listed in the following table:

Command	Description	CLI Mode
<code>switchport mode access</code>	Set the VLAN mode of the specified port to ACCESS mode. When this command is executed, the port is an untagged member of VLAN 1, and the port's default VLAN is 1.	Interface Configuration Mode
<code>switchport mode trunk</code>	Set the VLAN mode of the specified port to TRUNK mode. When this command is executed, the port is a tagged member of VLAN 1. The default VLAN for the port is 1.	Interface Configuration Mode
<code>no switchport trunk</code>	The VLAN mode of the port is no longer TRUNK mode and returns to the default, ACCESS mode.	Interface Configuration Mode
<code>switchport mode hybrid</code>	Set the VLAN mode of the specified port to HYBRID mode. When this command is executed, the port is an untagged member of VLAN 1, and the port's default VLAN is 1.	Interface Configuration Mode
<code>no switchport hybrid</code>	The VLAN mode of the port is no longer HYBRID mode and returns to the default, ACCESS mode.	Interface Configuration Mode

7.2.3 VLAN Configuration in ACCESS Mode

Specify the VLAN mode of the port as ACCESS mode before VLAN configuration. In this VLAN mode, the port defaults to an untagged member of VLAN 1, and the default VLAN for the port is 1.

The VLAN configuration commands in ACCESS mode are shown in the following table:

Command	Description	CLI Mode
<code>switchport access vlan <vlan-id></code>	Configure the port as an untagged member of the specified VLAN; the default VLAN for the port is the specified VLAN. The parameter ranges from 2 to 4094.	Interface Configuration Mode
<code>no switchport access vlan</code>	The port's VLAN configuration returns to the default, where the port is an untagged member of VLAN 1 and the port's default VLAN is 1.	Interface Configuration Mode

7.2.4 VLAN Configuration in TRUNK Mode

Specify the VLAN mode of the port as TRUNK mode before VLAN configuration. In this VLAN mode, the port defaults to a tagged member of VLAN 1, and the port's default VLAN is 1. The VLAN configuration commands in TRUNK mode are shown in the following table:

Command	Description	CLI Mode
<code>switchport trunk native vlan <vlan-id></code>	Configure the port's default VLAN, also known as the pvid. The parameter ranges from 2 to 4094.	Interface Configuration Mode
<code>switchport trunk allowed vlan all</code>	Configure the port as a tagged member of all VLANs; for new VLANs created later, the port is also a tagged member of those VLANs.	Interface Configuration Mode
<code>switchport trunk allowed vlan none</code>	The port is no longer a tagged member of all VLANs except VLAN 1.	Interface Configuration Mode
<code>switchport trunk allowed vlan add <vlan-list></code>	Configures the port as a tagged member of one or more specified VLANs. The <vlan-list> parameter can be a single VLAN, a VLAN range, or multiple VLANs, for example "1", "2-4", or "1,3,5".	Interface Configuration Mode
<code>switchport trunk allowed vlan remove <vlan-list></code>	Removes the port from one or more specified VLANs and removes its tagged membership of those VLANs. The <vlan-list> parameter can be a single VLAN, a VLAN range, or multiple VLANs, for example "1", "2-4", or "1,3,5".	Interface Configuration Mode

7.2.5 VLAN Configuration in HYBRID Mode

Set the port's VLAN mode to HYBRID before you configure VLANs on the port. In this VLAN mode, the port is an untagged member of VLAN 1 by default, and the port's default VLAN is 1. The VLAN configuration commands in HYBRID mode are shown in the following table:

Command	Description	CLI Mode
<code>switchport hybrid native vlan <vlan-id></code>	Configures the port as an untagged member of the specified VLAN and sets the port's default VLAN to that VLAN. The parameter range is 2 to 4094.	Interface Configuration Mode
<code>no switchport hybrid native vlan</code>	Removes the port from its default VLAN, so it is no longer a tagged or untagged member of that VLAN, and returns the port's default VLAN to 1.	Interface Configuration Mode
<code>switchport hybrid allowed vlan all</code>	Configures the port as a tagged member of all VLANs except VLAN1, and of VLANs created later.	Interface Configuration Mode
<code>switchport hybrid allowed vlan none</code>	The port is no longer a tagged or untagged member of all VLANs except VLAN 1, and the port's default VLAN returns to 1.	Interface Configuration Mode
<code>switchport hybrid allowed vlan add <vlan-list> egress-tagged enable</code>	Configures the port as a tagged member of one or more specified VLANs. The <vlan-list> parameter can be a single VLAN, a VLAN range, or multiple VLANs, for example "1", "2-4", or "1,3,5".	Interface Configuration Mode
<code>switchport hybrid allowed vlan add <vlan-list> egress-tagged disable</code>	Configures the port as an untagged member of one or more specified VLANs. The <vlan-list> parameter can be a single VLAN, a VLAN range, or multiple VLANs, for example "1", "2-4", or "1,3,5".	Interface Configuration Mode
<code>switchport hybrid allowed vlan remove <vlan-list></code>	Removes the port from one or more specified VLANs and removes its tagged or untagged membership of those VLANs. If the port's default VLAN is one of the specified VLANs, the default VLAN returns to 1.	Interface Configuration Mode

7.2.6 Viewing VLAN Information

The commands for viewing VLAN information are shown in the following table:

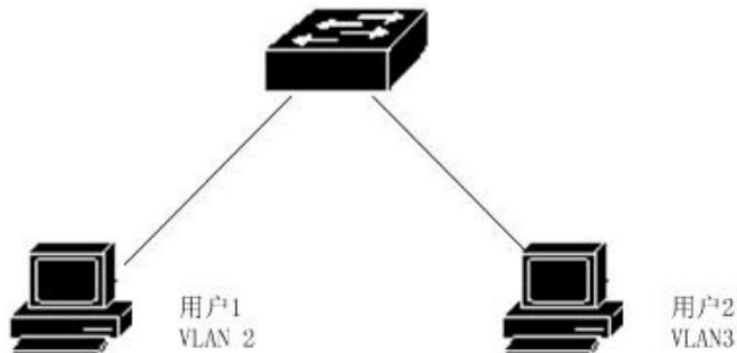
Command	Description	CLI Mode
<code>show vlan [vlan-id]</code>	If you enter no parameter, all VLAN information is displayed. If you enter a parameter, the information for the specified VLAN is displayed. The parameter range is 1 to 4094.	Normal Mode, Privileged Mode
<code>show interface</code>	Displays VLAN-related information for all ports on the system, such as VLAN mode and default VLAN.	Normal Mode, Privileged Mode
<code>show running-config</code>	Displays the current system configuration, so you can check the VLAN configuration.	Privileged Mode

7.3 VLAN Configuration Examples

7.3.1 PORT-Based VLAN

1) Configuration

There are two users, User 1 and User 2. They must be in different VLANs because they use different network functions and environments. User 1 belongs to VLAN2 and connects to port ge1/1 on the switch. User 2 belongs to VLAN3 and connects to port ge1/2 on the switch.



The switch is configured as follows:

Create a VLAN

```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
```

Assign a port to a VLAN

```
Switch#config t
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode access
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#exit
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 3
```

2) Debug

After configuration, you find that PCs in different VLANs cannot communicate. This is normal, because traffic between VLANs must be forwarded by Layer 3 routing. If PCs in the same VLAN cannot communicate, verify the following:

```
show vlan
```

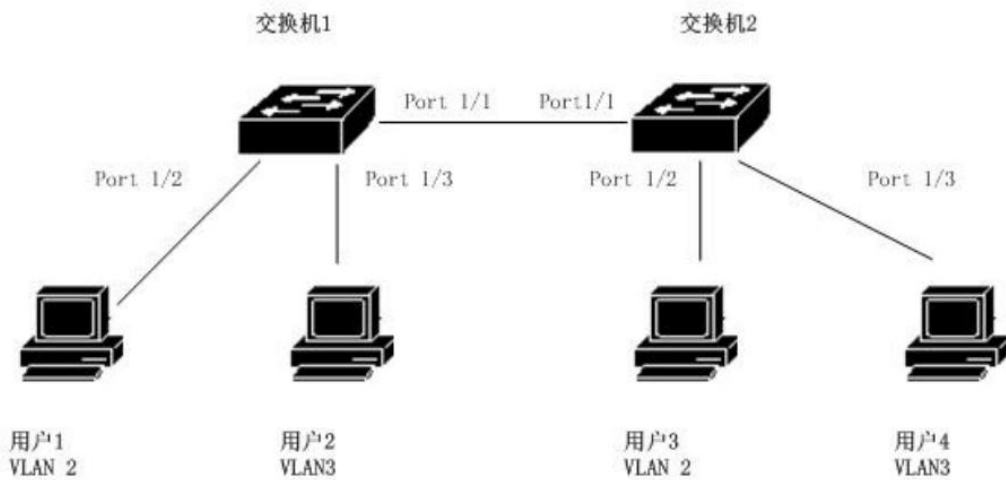
View the member ports of all VLANs

```
show vlan <vlan-id>
```

Check whether the port connected to a specific PC is in the specified VLAN

7.3.2 802.1Q-Based VLAN

1) Configuration



There are two switches connected to two users:

User	VLAN	Connected Port	Switch	Cascade Port
User1	2	1/2	switch1	1/1
User2	3	1/3	switch1	1/1
User3	2	1/2	switch2	1/1
User4	3	1/3	switch2	1/1

You must configure both switches.

Switch 1 Configuration:


```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
Switch#config t
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode trunk
Switch(config-ge1/1)#switchport trunk allowed vlan add 2
Switch(config-ge1/1)#switchport trunk allowed vlan add 3
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 2
Switch(config-ge1/2)#exit
Switch(config)#interface ge1/3
Switch(config-ge1/3)#switchport mode access
Switch(config-ge1/3)#switchport access vlan 3
```

Switch 2 Configuration:

```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
Switch#config t
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode trunk
Switch(config-ge1/1)#switchport trunk allowed vlan add 2
Switch(config-ge1/1)#switchport trunk allowed vlan add 3
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 2
Switch(config-ge1/2)#exit
Switch(config)#interface ge1/3
Switch(config-ge1/3)#switchport mode access
Switch(config-ge1/3)#switchport access vlan 3
```

2) Debug

The VLANs span both switches, so PCs in the same VLAN can communicate. If they cannot, check the following:

- Whether the port connected to the PC belongs to the corresponding VLAN and uses ACCESS mode to join that VLAN.
- Whether cascade port 1/1 is added to each VLAN and is in TRUNK mode.

7.4 MAC, IP Subnet, and Protocol VLANs

7.4.1 Introduction to MAC, IP Subnet, and Protocol VLANs

MAC-based VLANs are divided by the source MAC address of the packet. When the device receives an untagged packet (or a packet with tag 0) on a port, it determines the VLAN the packet belongs to from the source MAC address, and then automatically assigns the packet to the specified VLAN for transmission.

IP subnet-based VLANs are divided by the source IP address and subnet mask of the packet. When the device receives an untagged packet on a port, it determines the VLAN the packet belongs to from the source address, and then automatically assigns the packet to the specified VLAN for transmission. This feature is mainly used to transmit packets sent by a specified network segment or IP address in a specified VLAN.

A protocol-based VLAN assigns different VLAN IDs to the packets a port receives according to the protocol the packets belong to. The protocols that can be used to divide VLANs include IP, IPV6, and IPX.

7.4.2 MAC, IP Subnet, and Protocol VLAN Configuration

Before you configure VLANs based on MAC addresses, IP subnets, or protocols, you must create the corresponding VLANs.

Command	Description	CLI Mode
<code>mac-vlan mac WORD vlan <1-4094></code>	Creates a VLAN based on the source MAC address	VLAN Configuration Mode
<code>no mac-vlan mac WORD</code>	Deletes a VLAN based on the source MAC address	VLAN Configuration Mode
<code>no mac-vlan</code>	Deletes all VLANs based on the source MAC address	VLAN Configuration Mode
<code>mac-vlan enable</code>	Enables the MAC-VLAN feature on the interface	Interface Configuration Mode
<code>mac-vlan disable</code>	Disables the MAC-VLAN feature on the interface	Interface Configuration Mode
<code>show mac-vlan</code>	Displays all VLANs based on the source MAC address	Privileged Mode
<code>ip-subnet-vlan ip A.B.C.D A.B.C.D vlan <1-4094></code>	Creates a VLAN based on the source IP subnet	VLAN Configuration Mode
<code>no ip-subnet-vlan ip A.B.C.D A.B.C.D</code>	Deletes a VLAN based on the source IP subnet	VLAN Configuration Mode
<code>no ip-subnet-vlan</code>	Deletes all VLANs based on the source IP subnet	VLAN Configuration Mode
<code>ip-subnet-vlan enable</code>	Enables the interface's VLAN function based on the source IP subnet	Interface Configuration Mode
<code>ip-subnet-vlan disable</code>	Disables the interface's VLAN function based on the source IP subnet	Interface Configuration Mode
<code>show ip-subnet-vlan</code>	Displays all VLANs based on the source IP subnet	Privileged Mode
<code>protocol-vlan ether-type (ip ipv6 ipx ... <0-65535>) vlan <1-4094></code>	Creates a protocol-based VLAN	Interface Configuration Mode
<code>no protocol-vlan ether-type (ip ipv6 ipx ... <0-65535>)</code>	Deletes a protocol-based VLAN	Interface Configuration Mode

Command	Description	CLI Mode
<code>no protocol-vlan</code>	Deletes all protocol-based VLANs	Interface Configuration Mode
<code>show protocol-vlan</code>	Displays all protocol-based VLANs	Privileged Mode
<code>show vlan-partition interface IFNAME</code>	Shows the MAC and IP subnet-based VLANs enabled on an interface	Privileged Mode

7.5 Voice VLAN

7.5.1 Introduction to Voice VLAN

A Voice VLAN is a VLAN dedicated to users' voice data traffic. By creating a Voice VLAN and adding the port connected to the voice device to it, you can configure QoS (Quality of Service) parameters for the voice data to raise the priority of voice packets and ensure call quality.

The device determines whether a data stream is a voice data stream from the OUI field of the source MAC address in the packets entering the port. A packet whose source MAC address matches a voice equipment OUI address set on the system is treated as voice data traffic and is assigned to the Voice VLAN for transmission.

You can preset the OUI address or use the default OUI addresses as the matching criterion, as follows:

Serial Number OUI Address Manufacturer

1 0001-e300-0000 Siemens phone

2 0003-6b00-0000 Cisco phone

3 0004-0d00-0000 Avaya phone

4 00d0-1e00-0000 Pingtel phone

5 0060-b900-0000 Philips/NEC phone

6 00e0-7500-0000 Polycom phone

7 00e0-bb00-0000 3com phone

Manually add the IP phone access port to the Voice VLAN. The OUI address is then matched by identifying the source MAC address of the packet. After a successful match, the system issues ACL rules and configures the priority of the packet.

7.5.2 Voice VLAN Configuration

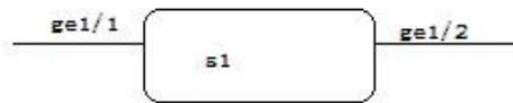
Before you configure Voice VLANs, you must create the corresponding VLANs.

Command	Description	CLI Mode
<code>voice-vlan oui WORD mask WORD</code>	Configures a user OUI	Global Configuration Mode

Command	Description	CLI Mode
<code>voice-vlan oui WORD mask WORD description WORD</code>	Configures a user OUI and names it	Global Configuration Mode
<code>no voice-vlan oui WORD mask WORD</code>	Deletes a user OUI configuration by OUI address and mask	Global Configuration Mode
<code>no voice-vlan oui description WORD</code>	Deletes a user OUI configuration by name	Global Configuration Mode
<code>no voice-vlan oui</code>	Deletes all user OUI configurations	Global Configuration Mode
<code>no voice-vlan default-oui WORD mask WORD</code>	Deletes a default OUI configuration by OUI address and mask	Global Configuration Mode
<code>no voice-vlan default-oui description WORD</code>	Deletes a default OUI configuration by name	Global Configuration Mode
<code>no voice-vlan default-oui</code>	Deletes all default OUI configurations	Global Configuration Mode
<code>voice-vlan default-oui resume</code>	Restores all default OUI configurations	Global Configuration Mode
<code>show voice-vlan oui</code>	Displays all default and user OUI configurations	Privileged Mode
<code>voice vlan <1-4094> (enable disable)</code>	Enables Voice VLAN on the interface	Interface Configuration Mode
<code>voice vlan qos remark cos <0-7> dscp <0-63></code>	Configures QoS priority on the interface, with a cos value of 6 and a dscp value of 46	Interface Configuration Mode
<code>no voice vlan qos</code>	Restores the default QoS priority configuration on the interface	Interface Configuration Mode
<code>no voice vlan</code>	Deletes the Voice VLAN configuration on an interface	Interface Configuration Mode
<code>show voice-vlan</code>	Shows how Voice VLANs are configured on all interfaces	Privileged Mode

7.5.3 Voice VLAN Configuration Example

Configure the voice data traffic (0009.ca00.0000) to enter through port ge1/1 and leave through port ge1/2 with tag 2, as shown in the following figure:



Configure switch S1 as follows:

```
Switch#con t
Switch(config)#vlan da
Switch(config-vlan)#vlan 2
Switch(config-vlan)#exit
Switch(config)#int ge1/1
Switch(config-ge1/1)#sw mod hy
Switch(config-ge1/1)#sw hybrid allowed vlan add 2 egress-tagged
disable
Switch(config-ge1/1)#voice vlan 2 en
Switch(config-ge1/1)#voice vlan 2 enable
Switch(config-ge1/1)#int ge1/2
Switch(config-ge1/2)#sw mod tr
Switch(config-ge1/2)#sw trunk allowed vlan add 2
Switch(config-ge1/2)#exit
Switch(config)#voice-vlan oui 0009.ca00.0000 mask ffff.ff00.0000
Switch(config)#
```

7.6 VLAN Mapping

7.6.1 Introduction to VLAN Mapping

The VLAN Mapping function can modify the VLAN Tag carried by a packet and provides the following mapping relationship: 1:1 VLAN Mapping, which changes the VLAN ID in the VLAN Tag carried by the packet to another VLAN ID. Before you can configure VLAN mappings, you must create the corresponding VLANs.

7.6.2 VLAN Mapping Configuration

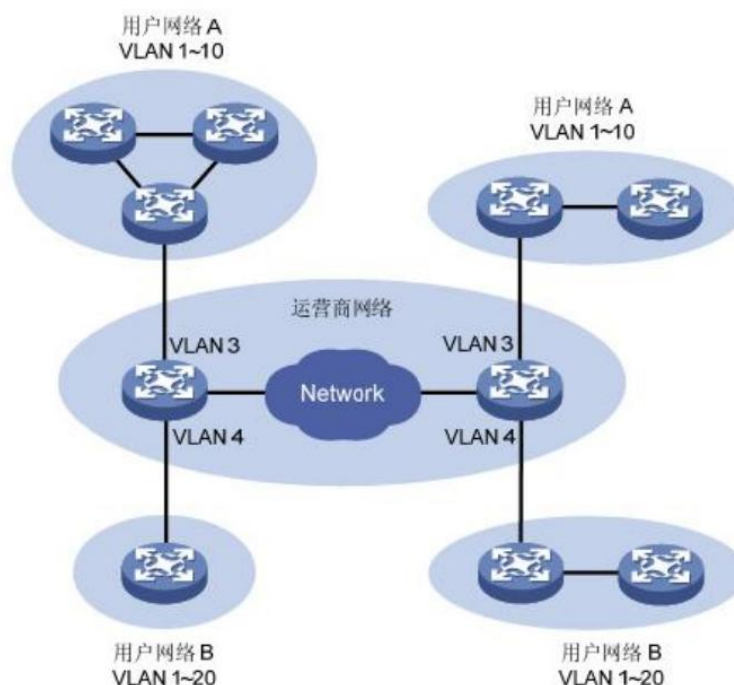
Command	Description	CLI Mode
<code>vlan-mapping vlan <1-4094></code> <code>map-vlan <1-4094></code>	Configures a VLAN mapping for the port	Interface Configuration Mode
<code>no vlan-mapping vlan <1-4094></code>	Deletes a VLAN mapping for a port	Interface Configuration Mode
<code>no vlan-mapping</code>	Deletes all VLAN mappings for a port	Interface Configuration Mode
<code>vlan-mapping enable</code>	Enables VLAN mapping on the port	Interface Configuration Mode
<code>vlan-mapping disable</code>	Disables VLAN mapping on the port	Interface Configuration Mode
<code>show vlan-mapping</code>	Displays all configured VLAN mappings	Privileged Mode

7.7 QinQ

7.7.1 Introduction to QinQ

The port QinQ feature provided by the device is a simple and flexible Layer 2 VPN technology. It encapsulates an outer VLAN Tag on the user's private network packet at the operator's network edge device, so the packet carries two VLAN Tags across the operator's backbone network (the public network). In the public network, the device forwards the packet according to the outer VLAN Tag only, and learns the packet's source MAC address into the MAC address table of the VLAN of the outer Tag. The user's private network VLAN Tag is carried as part of the packet data during transmission.

The QinQ feature lets operators use one VLAN to serve a user network that contains multiple VLANs. As shown in the figure below, the private VLANs of user network A are VLANs 1 to 10, and the private VLANs of user network B are VLANs 1 to 20. The operator assigns VLAN 3 to user network A and VLAN 4 to user network B. When a packet with the VLAN Tag of user network A enters the operator network, the packet is encapsulated with an outer VLAN Tag whose VLAN ID is 3. When a packet with the VLAN Tag of user network B enters the operator network, the packet is encapsulated with an outer VLAN Tag whose VLAN ID is 4. In this way, the packets of the different user networks stay completely separate while they cross the public network, and no confusion occurs even if the VLAN ranges of the two user networks overlap.



The Qinq feature enables the network to provide up to 4094X4094 VLANs, which meets the VLAN count requirements of a metropolitan area network. It mainly solves the following problems:

- (1) It eases the shortage of public network VLAN ID resources.
- (2) Users can plan their own private network VLAN IDs without conflicting with public network VLAN IDs.
- (3) It provides a simple Layer 2 VPN solution for a small MAN or enterprise network.

Qinq can be divided into two types: basic Qinq and flexible Qinq.

(1) Basic Qinq: Basic Qinq is implemented per port. After you enable basic Qinq on a port, the device tags each packet the port receives with the VLAN Tag of the port's default VLAN. A packet that already has a VLAN Tag becomes a double-tagged packet; a packet without a VLAN Tag becomes a packet with the port's default VLAN Tag.

(2) Flexible Qinq: Flexible Qinq is a more flexible implementation of Qinq, based on the combination of port and VLAN. In addition to all basic Qinq functions, it can take different actions on packets received on the same port according to their VLAN, and add different outer VLAN Tags to packets with different inner VLAN IDs.

7.7.2 Qinq Configuration

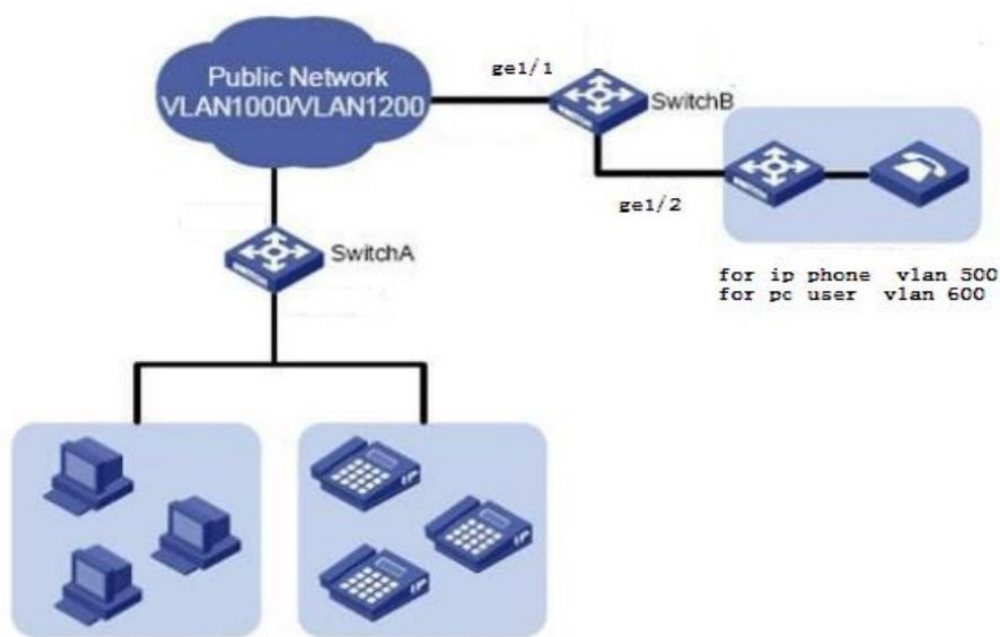
Command	Description	CLI Mode
<code>qinq tpid WORD</code>	Configures the tpid value carried in the port VLAN Tag, which is 0x8100 by default	Interface Configuration Mode
<code>no qinq tpid</code>	Restores the port's default tpid	Interface Configuration Mode

Command	Description	CLI Mode
qinq uplink	Configures the port as an uplink port	Interface Configuration Mode
no qinq uplink	Removes the uplink configuration from the port	Interface Configuration Mode
qinq customer	Configures the port as a customer port	Interface Configuration Mode
no qinq customer	Removes the customer configuration from the port	Interface Configuration Mode
qinq outer-vid <1-4094> inner-vid VLAN_ID	Configures one VLAN translation for the interface	Interface Configuration Mode
no qinq outer-vid <1-4094> [inner-vidVLAN_ID]	Deletes a VLAN translation for an interface	Interface Configuration Mode
show qinq	Displays the qinq status for all configurations	Privileged Mode

7.7.3 QinQ Configuration Examples

On SwitchB, port ge1/1 connects to the public network and port ge1/2 connects to the PC and the telephony server. The PC uses VLAN 600 and the IP phone uses VLAN 500. VLAN 100 and VLAN 200 are allowed to pass through the public network. The PC user data is transmitted across the public network in VLAN 200, and the IP phone data is transmitted across the public network in VLAN 100.

The networking diagram is as follows:



Configure Switch B as follows:

```
Switch#con t
Switch(config)#vlan da
Switch(config-vlan)#vlan 100
Switch(config-vlan)#vlan 200
Switch(config-vlan)#exit
Switch(config)#int ge1/1
Switch(config-ge1/1)#sw mod tr
Switch(config-ge1/1)#switchport trunk allowed vlan add 100
Switch(config-ge1/1)#switchport trunk allowed vlan add 200
Switch(config-ge1/1)#qinq uplink
Switch(config-ge1/1)#int ge1/2
Switch(config-ge1/2)#switchport mode hybrid
Switch(config-ge1/2)#switchport hybrid allowed vlan add 100 egress-tagged
```

dis

```
Switch(config-ge1/2)#switchport hybrid allowed vlan add 200 egress-tagged
```

dis

```
Switch(config-ge1/2)#qinq customer
Switch(config-ge1/2)#qinq outer-vid 100 inner-vid 500
Switch(config-ge1/2)#qinq outer-vid 200 inner-vid 600
Switch(config-ge1/2)#
Switch#show qinq
```

ifname tpid dtag-mode outer-vid inner-vid ge1/1 0x8100 uplink - -ge1/2 0x8100 customer 100 500 ge1/2 0x8100 customer 200 600

```
Switch#
```

Chapter 8 Configuring QoS

This chapter describes QoS and its configuration, and covers the following topics:

- QoS Introduction
- QoS Configuration
- QoS Configuration Example

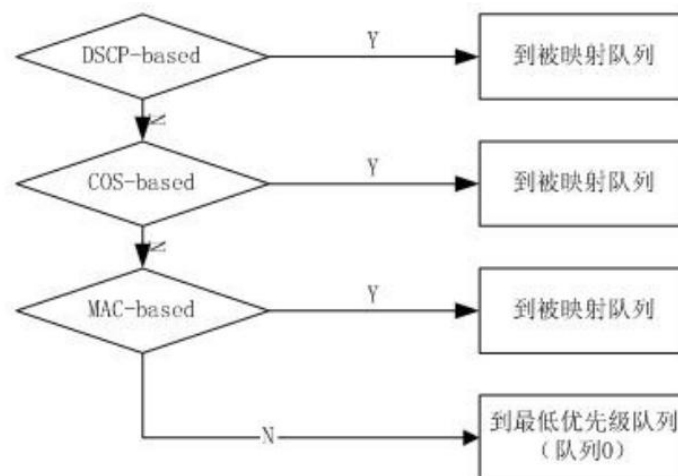
8.1 QoS Introduction

Use the QoS features of the switch to prioritize the important traffic forwarded through the switch. This makes better use of your network's bandwidth and makes network performance predictable.

In a switch, the priority information of a packet at an input port determines the queue the packet uses at an output port.

The switch implements QoS based on COS (802.1p), QoS based on DSCP (DiffServ), and QoS based on MAC. You can configure DSCP-based QoS on a physical port. COS-based QoS is enabled by default on all physical ports. MAC-based QoS is configured in the MAC binding function. You can configure only one QoS feature per physical port.

The following figure shows the packet forwarding process with QoS enabled:



The switch supports eight priority queues, 0 to 7, where queue 7 has the highest priority and queue 0 the lowest. The priority queues have four scheduling modes: SP, RR, WRR, and WDRR. SP is strict priority scheduling: it always forwards the packets of queue 7 first, does not start forwarding the packets of queue 6 until queue 7 is forwarded, does not start forwarding the packets of queue 5 until queue 6 is forwarded, and forwards the packets of queue 0 last. RR is round-robin scheduling: when the switch forwards packets, it polls the queues from the highest priority to the lowest, and each queue forwards one packet. WRR is polling with

priority: when forwarding packets, the switch polls the queues from the highest priority to the lowest according to the configured priorities. It first forwards the number of packets given by the priority from the highest priority queue, then the number of packets given by the priority from the next highest priority queue, and so on until the lowest priority queue has finished forwarding, after which it starts again from the highest priority queue. WDRR is weighted round-robin scheduling: for example, the weight of queue 3 is 4, so in one round it can forward 5 packets and in the next round it can forward only 3 packets. This flexibility suits highly aggregated network environments better.

To make configuration easier, we introduce the concept of the QosProfile. A QosProfile is an attribute that sets the mapping between 802.1p and the priority queue. You cannot configure this attribute. The mapping is shown in the following table:

QosProfile	802.1p (CoS) Value	Priority Queue
Qp0	0	0
Qp1	1	1
Qp2	2	2
Qp3	3	3
Qp4	4	4
Qp5	5	5
Qp6	6	6
Qp7	7	7

8.1.1 QoS Based on COS

By default, QoS over COS is enabled on the port. The switch reads the priority value of the VLAN TAG in the packet entering the port, and determines the output queue for the packet from the mapping between the configured COS value and the queue. If the packet has no VLAN TAG, or the VID of the VLAN TAG is 0, the switch fills in the packet according to the port's configured default VID and default priority, and then determines the output queue for the packet from that default priority.

8.1.2 DSCP-Based QoS

If DSCP-based QoS is enabled on a port, the switch reads the DSCP value in the IP packet entering the port and determines the output queue for the packet from the mapping between the configured DSCP value and the queue.

8.1.3 MAC-Based QoS

When a packet enters the switch, the switch searches its Layer 2 forwarding table using the destination MAC address of the packet and the VID of the packet's VLAN TAG. If a matching entry is found, the output queue for the packet is determined from the queue mapping configured in that entry.

8.1.4 Policy-Based QoS

QoS policies consist of classes, policies, and actions. Classes identify flows, and you can define a series of rules through commands to classify packets. Policy actions define the QoS actions taken on packets that match the rules. If policy-based QoS is enabled on a port, the switch classifies the packets entering the port. Packets that meet the classification requirements are processed according to the corresponding policy action. Packets that do not meet them are not processed, and their output queue is determined from the priority mapping.

8.2 QoS Configuration

8.2.1 Default Configuration for QoS

Configuration Item	Value	Configurable
Number of queues	8	No
Scheduling mode	WRR	Yes
Whether SP scheduling is enabled	disable	Yes
Whether RR scheduling is enabled	disable	Yes
Whether WDRR scheduling is enabled	disable	Yes
Queue weight	QP0[1],QP1[2],QP2[4],QP3[8] QP4[16],QP5[32],QP6[64] QP7[127]	Yes
Mapping between COS and qosprofile	COS0~ [qp0] COS1~ [qp1] COS2~ [qp2] COS3~ [qp3] COS4~ [qp4] COS5~ [qp5] COS6~ [qp6] COS7~ [qp7]	No
Mapping between DSCP and qosprofile	DSCP0~DSCP7[qp0] DSCP8~DSCP15[qp1] DSCP16~DSCP23[qp2] DSCP24~DSCP31[qp3] DSCP32~DSCP39[qp4] DSCP40~DSCP47[qp5] DSCP48~DSCP55[qp6] DSCP56~DSCP63[qp7]	Yes
Whether dscp-based QoS is enabled on the interface	disable	Yes
Whether COS-based QoS is enabled on the interface	enable	No
Interface user priority (COS value)	0	Yes

8.2.2 Configuring the Scheduling Mode

The default scheduling mode of the switch is WRR. You can configure the SP, RR, and WDRR scheduling modes with commands.

Command	Description	CLI Mode
<code>qos sched {rr sp wrr wdr}</code>	Configures the QoS scheduling mode	Interface Configuration Mode

8.2.3 Configuring Queue Weight

Command	Description	CLI Mode
<code>qos qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7} weight<1-127></code>	Configures the weight of each priority queue	Interface Configuration Mode
<code>no qos qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7} weight</code>	Restores the queue weight configuration to the default	Interface Configuration Mode

The queue weight is the number of packets a priority queue forwards in one round of polling. When you configure the queue weight, make sure the weight of a low priority queue does not exceed the weight of a high priority queue.

8.2.4 Configure the Mapping Between DSCP and QosProfile

Command	Description	CLI Mode
<code>qos dsc-map-qp <0-63> qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7}</code>	Configures the mapping between DSCP and QosProfile	Interface Configuration Mode
<code>no qos qosprofile {qp0 qp1 qp2 qp3 qp4 qp5 qp6 qp7} weight</code>	Restores the mapping between DSCP and qosprofile to the default configuration.	Interface Configuration Mode

8.2.5 Configuring Port DSCP-Based QoS

You can configure QoS functions only on physical ports, not on TRUNK or Layer 3 interfaces.

Command	Description	CLI Mode
<code>qos dscp-based</code>	Enables the port's DSCP-based QoS function.	Interface Configuration Mode
<code>no qos dscp-based</code>	Disables the port's DSCP-based QoS function.	Interface Configuration Mode

8.2.6 Configuring Port User Priority (COS Value)

Command	Description	CLI Mode
<code>qos user-priority <0-7></code>	Configures the user priority of the port (COS value)	Interface Configuration Mode
<code>no qos user-priority</code>	Restores the port's user priority (COS value) to the default configuration.	Interface Configuration Mode

8.3 QoS Configuration Example

Configure the user priority (COS value) of ge1/3 as 3. COS-based QoS is enabled by default:

```
Switch#configure terminal
Switch#(config)#interface ge1/3
Switch#(config-ge1/3)#qos user-priority 3
Switch#(config-ge1/3)#end
```

Configure interface ge1/3 to enable DSCP-based QoS, with DSCP value 3 mapped to priority queue 2:

```
Switch#configure terminal
Switch#(config)#interface ge1/3
Switch#(config-ge1/3)#qos dscp-map-qp 3 qosprofile qp2
Switch#(config-ge1/3)#qos dscp-based
Switch#(config-ge1/3)#end
```

8.4 Policy QoS Configuration Example

Configure ACLs to capture the traffic from source MAC1, MAC2, and MAC3

(You can modify the ACL rules as required. This is only a simple example.)

```
access-list 700 permit host 0000.0000.1111 vid any ip any any
access-list 701 permit host 0000.0000.2222 vid any ip any any
access-list 702 permit host 0000.0000.3333 vid any ip any any
```

Configure the QoS classes to match the traffic from source MAC1, MAC2, and MAC3 respectively

(You can modify the matching rule cos or dscp as required. This is only a simple example.)

```
qos class 10 match acl 700 qos class 11 match acl 701 qos class 12 match acl 702
```

Configure the QoS policy to re-mark the 802.1p priority of the traffic from MAC1, MAC2, and MAC3 respectively

(You can modify the policy as required. This is only a simple example.)

qos policy 10 class 10 remark cos 7 qos policy 10 class 11 remark cos 5 qos policy 10 class 12 remark cos 3

Apply the QOS policy to the port

```
interface ge1/2
qos apply-policy 10
```

Chapter 9 Configuring MSTP

This chapter describes MSTP and its configuration, and covers the following topics:

- MSTP Introduction
- MSTP Configuration
- MSTP Configuration Example

9.1 MSTP Introduction

The switch supports the IEEE 802.1d, IEEE 802.1w, and IEEE 802.1s STP protocols.

9.1.1 Overview

MSTP uses RSTP to converge multiple VLANs quickly into a spanning-tree instance, and each instance has a spanning-tree topology that is independent of the other spanning-tree instances. This architecture provides multiple forwarding paths for traffic, enables load balancing, and reduces the number of spanning tree instances needed to support a large number of VLANs.

9.1.2 Multiple Spanning Tree Domain

For instances to take part in Multiple Spanning Tree (MST) calculations, the switches must be configured consistently with the same MST configuration information. A group of connected switches with the same MST configuration forms an MST domain.

The MST configuration determines the domain each switch belongs to. The configuration includes the domain name, the revision number, and the MST instance and VLAN assignment mappings. This information generates a unique Digest in the MST configuration. The digests within the same domain are the same and must be the same, as you can see with the `show spanning-tree MST config` command.

A domain can have one or more members with the same MST configuration, and each member must be able to handle RSTP BPDUs. There is no limit to the number of MST domains in a network, but each domain supports up to 16 instances. You can assign only one VLAN to a spanning-tree instance at a time.

9.1.3 IST, CIST, and CST

Internal Spanning Tree (IST): a spanning tree that runs within the MST domain.

In each MSTP domain, MSTP maintains multiple spanning tree instances. Instance 0 is a special instance of a domain called the IST. All other MST instances are numbered 1 through 15.

The IST is simply a spanning tree instance that receives and sends BPDUs; the information for all other spanning tree instances is compressed into MSTI BPDUs. Because MSTI BPDUs carry the information for all instances, the number of BPDUs a switch that supports multiple spanning tree instances must process is reduced.

All MST instances in the same domain share the same protocol timer, but each MST instance has its own topology parameters, such as root switch ID and root path cost. By default, all VLANs are assigned to the IST.

The Common and Internal Spanning Tree (CIST) is the collection of all ISTs in each MST domain plus the Common Spanning Tree, which connects the MST domains into a single spanning tree.

A spanning tree computed within a domain looks like a subtree of the CST that contains all the switch domains. The CIST is formed by a spanning tree computation run between switches that support the 802.1W and 802.1D protocols. The CIST inside the MST domain is the same as the CST outside the domain.

Common Spanning Tree (CST): a spanning tree that runs between MST domains.

9.1.4 Intra-Domain Operation

The IST connects all MSTP switches in a domain. When the IST converges, the root of the IST becomes the IST master, which is the switch in the domain with the lowest bridge ID and path cost to the CST root. If there is only one domain in the network, the IST master is also the CST root. If the CST root is outside the domain, an MSTP switch at the boundary of the domain is selected as the IST master.

When an MSTP switch initializes, it sends BPDUs claiming itself as the CST root and IST master, with the path cost to the CST root and IST master set to 0. The switch also initializes all MST instances and claims to be their root. If the switch receives MST root information that is better than the information stored on the current port (a lower bridge ID, a lower path cost, and so on), it gives up its claim to be the IST master.

During initialization, a domain may have many subdomains, each with its own IST master. When a switch receives a higher priority IST message, it leaves its old subdomain and joins a new subdomain that may contain the true IST master. In this way all subdomains contract, except the subdomain that contains the true IST master.

For correct operation, all switches in the MST domain must recognize the same IST master. Two switches in the domain therefore synchronize the port roles of one of their MST instances only if they converge to a common IST master.

9.1.5 Interdomain Operation

If the network contains multiple domains or legacy 802.1D switches, MSTP establishes and maintains the CST, which includes all the MST domains and all the legacy STP switches in the network. MST instances are associated with the IST at the domain boundary to become part of the CST.

The IST connects all switches within the MSTP domain and looks like a subtree of the CST, which encloses all switch domains. The root of the subtree becomes the IST master. The MST domain looks like a virtual switch adjacent to the STP switch and the MST domain.

Only the CST instance sends and receives BPDUs. The MST instances add their spanning-tree information to the BPDUs to interact with neighboring switches and compute the final spanning-tree topology. For this reason, the spanning tree parameters related to BPDU delivery, such as hello time, forward time, max-age, and max-hops, are configured only in the CST instance and do not affect all MST instances. Parameters related to the spanning tree topology, such as switch priority, port VLAN cost, and port VLAN priority, can be configured in both the CST instance and the MST instance.

MSTP switches use version 3 RSTP BPDUs or 802.1D BPDUs to communicate with 802.1D switches. MSTP switches use MSTP BPDUs to communicate with MSTP switches.

9.1.6 Hop Count

The IST and MST instances do not use the message-age and maximum-age information in the BPDUs that are configured to compute the spanning-tree topology. Instead, they use the path cost to the root and a hop-count mechanism equivalent to the IP TTL.

You can configure the maximum hop count for a domain and apply it to that domain's IST and all MST instances. The hop count works the same way as the message-age result, which is determined after a reconfiguration. The root switch of the instance always sends a BPDU (or M-record) with a cost of 0 and a hop-count of Max. When a switch receives a BPDU, it decrements the remaining hop count by one and propagates the remaining hop count in the BPDUs it generates. When the count reaches 0, the switch discards the BPDU and ages out the information for that port.

Within a domain, the message-age and maximum-age information in the RSTP BPDU stays consistent, and the same value is propagated across the designated ports of the boundary domain.

9.1.7 Boundary Port

A boundary port is a port that connects the MST domain to a spanning-tree domain running RSTP alone, to a spanning-tree domain running 801.1D alone, or to an MST domain with a different configuration. A boundary port is also a port connected to a LAN whose designated switch is either a single spanning-tree switch or a switch with a different MST domain configuration.

On boundary ports, the MST port role is not important, and the MST port state is forced to match the IST port state. When the IST port is forwarding, the MST port on the boundary is forwarding. An IST port at the boundary can have any role except backup port.

On a shared boundary connection, the MST port waits in the blocking state for the forward-delay time to expire before it transitions to the learning state. It then waits for another forward-delay time to expire before it transitions to forwarding.

If the boundary port is a point-to-point connection and is the IST root port, the MST port transitions to the forwarding state as soon as the IST port transitions to the forwarding state.

If a boundary port transitions to the forwarding state in one instance, it is forwarding in all instances, and a topology change is triggered. If a boundary port with an IST root or designated port role receives a topology change advertisement, the MSTP switch triggers a topology change on the IST instance and on all MST instances that are active on that port.

9.1.8 Interoperability of MSTP and 802.1d STP

A switch running MSTP supports a built-in protocol migration mechanism that lets it work with 802.1D. If the switch receives an 802.1D configuration BPDU on a port, it sends 802.1D BPDUs on that port. An MSTP switch can detect when an 802.1 BPDU is received on a boundary port of a domain, and when an MSTP BPDU or RSTP BPDU is received on a boundary port of a different domain.

However, if a switch stops receiving 802.1D BPDUs, it does not automatically revert to MSTP mode, because it cannot tell whether the other switch has been removed from the connection unless the other switch is the designated switch. Also, the switch may keep assigning the boundary port role to a port after the switch

connected to it has joined the domain. Restart the protocol migration process to force negotiation with the neighbor switch.

If all the other switches on the connection are RSTP switches, they can process MSTP BPDUs as well as RSTP BPDUs. The MSTP switch therefore sends either a version 0 configuration and TCN BPDU or a version 3 MSTP BPDU on the boundary port. A boundary port is connected to a LAN whose designated switch is either a single spanning-tree switch or a switch with a different MST configuration.

9.1.9 Port Role

MSTP adopts the fast convergence algorithm of RSTP. MSTP port roles and fast convergence are described briefly below together with RSTP.

RSTP provides fast convergence for assigning port roles and determining the active topology. RSTP is based on IEEE 802.1D STP, which chooses the switch with the highest priority as the root switch. RSTP assigns a port one of the following roles:

- Root port: provides the best path cost when the switch forwards packets to the root switch.
- Designated port: connects to a designated switch. It has the lowest path cost when forwarding packets from the LAN to the root switch. The port through which a designated switch connects to a LAN is called a designated port.
- Alternate port: provides an alternate path to the root switch for the current root port.
- Backup port: acts as a backup for the designated port on the path toward the leaves of the spanning tree. A backup port exists only when two ports are connected together in a point-to-point loop, or when a switch has two or more ports connected to a shared LAN segment.
- Disabled port: has no port role in the spanning-tree operation.
- Master port: located at the domain root or on the shortest path to the overall root, it is the port that connects the domain to the overall root.

The root port and designated port roles are included in the active topology. The alternate and backup port roles are not.

In a stable topology with fixed port roles throughout the network, RSTP ensures that every root port and designated port moves to the forwarding state immediately, while all alternate and backup ports stay in the discarding state. The port state controls forwarding and learning.

Fast Convergence

RSTP provides fast recovery from a switch failure, a port failure, or a LAN failure, and it provides fast recovery for edge ports, new root ports, and point-to-point connections:

- Edge ports: if you configure a port as an edge port, it transitions to the forwarding state immediately. Enable this on boundary ports only if the port is connected to a single terminal or to a device that is known not to need spanning tree computation.
- Edge ports: if you configure a port as an edge port, it transitions to the forwarding state immediately. Enable this on boundary ports only if the port is connected to a single terminal or to a device that is known not to need spanning tree computation.

- Root ports: if RSTP selects a new root port, it blocks the old root port and moves the new root port to the forwarding state immediately.

Topology Change

This section describes how RSTP and 802.1D differ in handling spanning-tree topology changes.

- Detection: in 802.1D, any transition between the blocking and forwarding states causes a topology change. In RSTP, only a transition from blocking to forwarding causes a topology change, because topology changes are considered only when they increase connectivity. A state change at an edge port does not cause a topology change. When an RSTP switch detects a topology change, it floods the information it has learned to all non-edge ports except the port that received the TC information.
- Notification: unlike 802.1D, RSTP does not use the TCN BPDUs. However, the RSTP switch processes and generates TCN BPDUs for interoperability with 802.1D. Acknowledgement: when an RSTP switch receives a TCN message on a designated port from an 802.1D switch, it responds with an 802.1D BPDU that has the TCA flag set. However, if the TC-while timer (the same topology-change timer as in 802.1D) is active, the switch is connected to the 802.1D switch on the root port, and it receives a configuration BPDU with a TCA, the TC-while timer is reset. This behavior is required only to support 802.1D switches.

RSTP BPDUs never have the TCA flag.

- Propagation: when an RSTP switch receives a TC message from another switch on a designated or root port, it propagates the message to all non-edge ports, designated ports, and root ports, except the receiving port. All these ports on the switch start a TC-while timer and flood the information they learn.
- Protocol migration: for backward compatibility with 802.1D switches, RSTP sends 802.1D configuration BPDUs and TCN BPDUs selectively on a per-port basis.

When a port has been initialized, the migrate-delay timer starts, which sets the minimum time during which RSTP BPDUs are sent, and RSTP BPDUs are sent. While the timer is active, the switch processes all BPDUs received on the port and ignores the protocol type.

If the switch receives an 802.1D BPDU after the port's migration-delay timer has stopped, it assumes it is connected to an 802.1D switch and starts using 802.1D protocol BPDUs. However, if the RSTP switch is using 802.1D BPDUs on a port and receives an RSTP BPDU after the timer stops, that port restarts the timer and begins using RSTP BPDUs.

9.1.10 Introduction to 802.1D Spanning Tree

The Spanning Tree Protocol is based on the following:

- 1) There is a unique group address (01-80-C2-00-00-00) that identifies all switches on a particular LAN. All switches recognize this group address;
- 2) Each switch has a unique identifier (Bridge Identifier);
- 3) Each switch port has a unique Port Identifier. Managing the spanning tree configuration also requires a relative priority for each switch, a relative priority for each port on each switch, and a path cost for each port.

The switch with the highest priority is called the root switch. Each switch port has a root path cost, which is the sum of the path costs of the network segments between the switch and the root switch. The port with the lowest root path cost on a switch is called the root port. If several ports have the same root path cost, the port with the highest priority is the root port.

In each LAN there is one switch, called the designated switch, which is the switch with the lowest root path cost in that LAN. The port that connects the LAN to the designated switch is the designated port of the LAN. If more than two ports on the designated switch are connected to this LAN, the port with the highest priority is selected as the designated port.

The decisions needed to form a spanning tree:

1) Decide on the root switch. a. Initially, all switches consider themselves to be the root switch. b. Each switch broadcasts configuration BPDUs to the LANs it is connected to, with the same values in root_ID and bridge_ID. c. When a switch receives a configuration BPDU from another switch, if the value of the root_ID field in the received configuration BPDU is greater than the value of the switch's own root_ID parameter, the frame is discarded. Otherwise, the switch updates its root_ID, root_path_cost, and other parameters, and continues to broadcast configuration BPDUs with the new values.

2) Determine the root port. a. The port on a switch that has the lowest root path cost is called the root port.

b. If several ports have the same lowest root path cost, the port with the highest priority is the root port. If two or more ports have the same lowest root path cost and the same highest priority, the port with the lowest port number is the root port by default.

3) Identify the designated switch for the LAN. a. Initially, all switches consider themselves the designated switch for the LAN.

b. When a switch receives a BPDU from another switch in the same LAN with a lower root path cost, the switch no longer claims to be the designated switch. If two or more switches in a LAN have the same root path cost, the switch with the highest priority is selected as the designated switch.

c. If at some point the designated switch receives a message from another switch on the LAN because of competition for the designated switch role, the designated switch sends a configuration BPDU back to determine the designated switch again.

4) Determine the designated port

The port on the LAN's designated switch that connects to the LAN is the designated port. If the designated switch has two or more ports connected to the LAN, the port with the lowest ID is the designated port.

All ports except the root port and the designated port are placed in the blocking state. Once the root switch, the root port of each switch, and the designated switch and designated port of each LAN are determined, the spanning-tree topology is determined.

9.2 MSTP Configuration

9.2.1 Default Configuration

Command Parameter	Default Value
spanning-tree mst enable (Enable MSTP)	shut
spanning-tree mst priority (Switch cist priority)	32768
spanning-tree mst hello-time (Switch cist hello-time)	2 seconds
spanning-tree mst forward-time (Switch cist forward-time)	15 seconds
spanning-tree mst max-age (Switch cist max-age)	20 seconds
spanning-tree mst max-hops (Switch cist max-hops)	20 seconds
instance 1 priority (Instance priority)	32768
spanning-tree mst instance 1 priority (Port instance priority)	128
spanning-tree mst instance 1 path-cost (Port instance path-cost)	20000000
spanning-tree mst priority (Port cist priority)	128
spanning-tree mst path-cost (Port cist path-cost)	20000000

9.2.2 General Configuration

Start MSTP

By default, MSTP is off when the system starts.

The configuration process to start MSTP is:

```
Switch#configure terminal
Switch(config)#spanning-tree mst enable
```

The command to turn off MSTP is:

```
Switch#configure terminal
Switch(config)#no spanning-tree mst
```

Configure max-age

The max-age configuration applies to all instances. max-age is the number of seconds the switch waits to receive spanning-tree configuration information before it triggers a reconfiguration.

The default is 20 seconds, and the configuration range is 6 to 40 seconds.

Configuration process:

```
Switch#configure terminal
Switch(config)#spanning-tree mst max-age <seconds>
```

Configure max-hops

max-hops is the number of hops allowed within a domain before a BPDU is dropped.

The default value is 20, and the configuration range is 1 to 40.

Configuration process:

```
Switch#configure terminal
Switch(config)#spanning-tree mst max-hops <hop-count>
```

Configure forward-time

The forward-time configuration applies to all instances. forward-time is the number of seconds the port waits to move from discarding to learning and from learning to forwarding.

The default is 15 seconds, and the configuration range is 4 to 30 seconds. According to the spanning tree protocol, forward-time must satisfy the following condition: $2 * (\text{forward-time} - 1) \geq \text{max-age}$. Configuration process:

```
Switch#configure terminal
Switch(config)#spanning-tree mst forward-time <seconds>
```

Configure hello-time

The hello-time configuration applies to all instances. hello-time is the interval at which the root switch generates configuration information.

The default is 2 seconds, and the configuration range is 1 to 10 seconds. According to the spanning tree protocol, hello-time must satisfy the following condition: $2 * (\text{hello-time} + 1) \leq \text{max-age}$.

Configuration process:

```
Switch#configure terminal
Switch(config)# spanning-tree mst hello-time <seconds>
```

Configure the priority of the CIST bridge

The default is 32768 and the configuration range is <0-61440>. The CIST priority value can only be a multiple of 4096.

Configuration process:

```
Switch#configure terminal
Switch(config)#spanning-tree mst priority <priority>
```

Configure CISCO compatibility

This switch uses the MSTP protocol based on 802.1s, in which each MSTI message is 16 bytes long, while each MSTI message in the BPDU of a CISCO switch is 26 bytes long. To interoperate with CISCO switches, enable CISCO compatibility when you configure the switches in this network.

When CISCO compatibility is enabled, two domains are considered the same as long as the domain name and the revision number are the same.

This feature is not enabled by default.

Enable CISCO compatibility:

```
Switch#configure terminal
Switch(config)#spanning-tree mst cisco-interoperability enable
```

Disable CISCO compatibility:

```
Switch#configure terminal
Switch(config)#spanning-tree mst cisco-interoperability disable
```

Reset the protocol detection task

For compatibility with the 802.1D STP protocol, the system can automatically detect the protocol running on the other system. It determines the protocol to run on this port from the protocol running on the other side.

In some cases the protocol must be reset. For example, a port negotiates to run the STP protocol, and after a while the peer device running STP is replaced by a host. You now need to configure the port as a fast port, but the port already runs the STP protocol and the protocol negotiation task has stopped. In this case, reset the protocol negotiation task so that the port renegotiates the protocol with the host.

Reset the protocol detection task for the entire device:


```
Switch#clear spanning-tree detected protocols
```

Reset the protocol detection task for a port:

```
Switch#clear spanning-tree detected protocols interface <if-name>
```

9.2.3 Domain Configuration

If two or more devices are in the same domain, they must have the same VLAN instance mapping, the same revision number, and the same domain name.

A domain has one or more members with the same MST configuration, and each member must be able to handle RSTP BPDUs. There is no limit on the number of members in a network, but each domain can support up to 16 instances.

Instance configuration is described in Instance Configuration. Only the domain name and revision number configuration are described here.

Configure the domain name:

```
Switch#configure terminal
Switch(config)#spanning-tree mst configuration
Switch(config-mst)#region <region-name>
```

Configure the revision number:

```
Switch#configure terminal
Switch(config)#spanning-tree mst configuration
Switch(config-mst)# revision <revision-num>
```

9.2.4 Instance Configuration

The system supports 16 instances, with instance ID numbers from 0 to 15. A VLAN can be assigned to only one spanning-tree instance at a time.

By default, there is only one instance, 0, and all VLANs belong to it.

The process of configuring an instance:

```
Switch#configure terminal
Switch(config)#spanning-tree mst configuration
Switch(config-mst)#instance <instance-id> vlan <vlan-id>
```

Configure the priority of the MSTI bridge

The default is 32768 and the configuration range is <0-61440>. The MSTI priority value can only be a multiple of 4096.

Configuration process:

```
Switch#configure terminal
Switch(config)#spanning-tree mst configuration
Switch(config-mst)#instance <instance-id> priority <priority>
```

9.2.5 Port Configuration

The following describes the port configuration related to MSTP. Only the simple configuration is described here; port fast and root guard are described separately later.

The process of configuring a port to join an instance:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst instance <instance-id>
```

Configure the priority of the CIST port

The default is 128 and the configuration range is <0-240>. The priority value of the CIST port can only be a multiple of 16.

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst priority <priority>
```

Configure the priority of the MSTI port

The default is 128 and the configuration range is <0-240>. The priority value of the MSTI port can only be a multiple of 16.

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst instance <instance-id> priority
```

<priority>

Configure the path cost for a CIST port

The default is 20000000, and the configuration range is 1-200000000. The following table maps bandwidth to path cost:

Bandwidth	Path Cost
100,000(100K)	200000000
1,000,000(1M)	20000000
10,000,000(10M)	2000000
100,000,000(100M)	200000
1,000,000,000(1G)	20000
10,000,000,000(10G)	2000
100,000,000,000(100G)	200
1,000,000,000,000(1T)	20
>1000000000000	2

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst path <path-cost>
```

Configure the path-cost for an MSTI port

The default is 200000000, and the configuration range is 1-2000000000. Bandwidth and path costs are the same as in the table above.

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst instance <instance-id> path-cost
```

<path-cost>

Configure the version number of the protocol packets that are sent

By default, the switch sends MSTP protocol packets. The configuration range is 0-3, and the mapping is 0-stp, 2-rstp, and 3-mstp.

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)# spanning-tree mst force-version <version-id>
```

Configure the connection type

If a port connects to another port in a point-to-point manner and the local port becomes a designated port, RSTP determines a loop-free topology by negotiating a fast transition with the connected port through the proposal-agreement process.

The proposal-agreement negotiation process is described briefly below.

When a switch receives a proposal message on one of its ports and that port is selected as the new root port, RSTP forces all other ports to synchronize with the new root port information.

The switch is synchronized when all other ports are synchronized with the superior root information received from the root port.

When RSTP forces a port to synchronize the new root information, a designated port that is in the forwarding state and is not configured as an edge port moves to the blocking state. Normally, when RSTP forces a port to synchronize new root information and the port cannot meet the above conditions, the port state is set to blocking.

After ensuring that all ports are synchronized, the switch sends an agreement message to the root port corresponding to the designated port. When the switches on a point-to-point connection agree on their port roles, RSTP immediately moves the port state to forwarding.

On a shared connection, the port state is determined by an 802.1D calculation.

The default port connection type is point-to-point.

Configure the port connection type as a point-to-point connection:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst link-type point-to-point
```

Configure the port connection type as a shared connection:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config-ge1/2)#spanning-tree mst link-type shared
```

9.2.6 PortFast Related Configuration

1) Port Fast

Port Fast immediately moves an access or trunk port from blocking to forwarding, bypassing the listening and learning states. Use Port Fast on a port that connects to a single workstation or server, so that the device joins the network immediately without waiting for the spanning tree to converge.

Configure a port as a fast port:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst portfast
```

2) BPDU Filtering

You can turn on BPDU filtering globally on the switch or on a per-port basis, but the two behave differently.

At the global level, use the spanning-tree MST portfast bpdu-filter command to enable BPDU filtering on ports in the portfast bpdu-filter default state.

At the port level, use spanning-tree MST portfast bpdu-filter enable to turn on the BPDU filter on any port.

This feature prevents a Port Fast port from receiving or sending BPDUs. Configure BPDU filtering:

In global configuration mode:

```
Switch#configure terminal
Switch(config)# spanning-tree mst portfast bpdu-filter
```

In interface configuration mode:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst portfast bpdu-filter enable
```

3) BPDU Guard

You can turn on the BPDU guard feature globally on the switch or on a per-port basis, but the two behave differently.

At the global level, use spanning-tree MST portfast bpdu-guard to turn on BPDU guard for ports in the portfast bpdu-guard default state.

At the port level, you can turn on BPDU guard on any port.

When a port configured with BPDU guard receives a BPDU, the spanning tree shuts down the port. In a valid configuration, Port Fast-enabled ports do not receive BPDUs. A BPDU received on a Port Fast-enabled port indicates an invalid configuration, such as the connection of an unauthorized device, and BPDU guard enters the error-disabled state.

Error-disabled means that when a port with BPDU guard receives a BPDU, the error-disable timer starts if the system is configured with the error-disable mechanism. Error-disable restarts the port after the system-configured timeout.

In global configuration mode:

```
Switch#configure terminal
Switch(config)# spanning-tree mst portfast bpdu-guard
```

In interface configuration mode:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst portfast bpdu-guard enable
```

Configure error-disable

Enable the error-disable mechanism

```
Switch#configure terminal
Switch(config)#spanning-tree mst errdisable-timeout enable
```

Configure the error-disable timeout

```
Switch#configure terminal
Switch(config)#spanning-tree mst errdisable-timeout interval <seconds>
```

9.2.7 Root Guard Related Configuration

An SP Layer 2 network can contain many connections to switches that are not its own. In such a topology, the spanning tree can reconfigure itself and select a customer switch as the root switch. You can avoid this by configuring root guard on the SP switch port that connects to the switch on the customer network. If the spanning tree calculation selects a port on the customer network as the root port, root guard sets that port to the root-inconsistent (blocked) state, which prevents the customer switch from becoming the root switch or having a path to the root.

If a switch outside the SP network becomes the root switch, the port is blocked (root-inconsistent state) and the spanning tree selects a new root switch. The customer switch does not become the root switch, and it has no path to the root.

If the switch operates in MST mode, root guard forces the port to be a designated port. If a boundary port is blocked because of root guard in the IST instance, the port is blocked in all MST instances. A boundary port is a port that is connected to a LAN. The designated switch is either an 802.1D switch or a switch with a different MST domain configuration.

Root guard applies to all VLANs to which a port belongs when that port is turned on. VLANs can be aggregated and mapped to an MST instance.

Configuration process:

```
Switch#configure terminal
Switch(config)#interface <if-name>
Switch(config)#spanning-tree mst guard root
```

9.3 MSTP Configuration Example

(1) Configuration:

Three switches are connected in a ring, and the spanning tree protocol must be turned on for each switch to avoid loops. Perform the configuration on each switch individually.

Switch 1 Configuration:

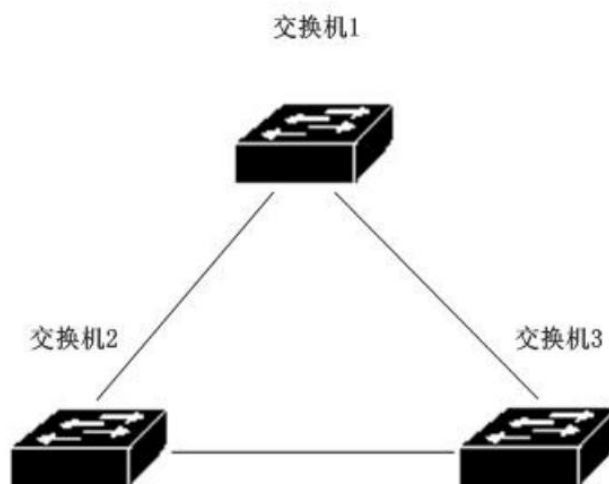
```
Switch>en
Switch#configure terminal
Switch(config)#spanning mst enable
```

Switch 2 Configuration:

```
Switch>en
Switch#configure terminal
Switch(config)#spanning mst enable
```

Switch 3 configuration:

```
Switch>en
Switch#configure terminal
Switch(config)#spanning mst enable
```



(2) Debugging:

To see which switch is selected as the root bridge:

Execute `show spanning-tree MST` and observe that the value of CISTRoot is the lowest MAC address of the three switches, which means the root election result is correct.

```
Switch#show spanning-tree mst
```

To view the port status of a switch in the spanning tree:

Execute the `show spanning-tree MST interface ge1/1` command to observe the State value of port ge1/1 in instance 0

```
Switch#show spanning-tree mst interface ge1/1
```


Chapter 10 Configuring EAPS

This chapter describes EAPS and its configuration, including the following topics:

- EAPS introduction
- EAPS basic concepts
- EAPS protocol introduction
- EAPS configuration
- Restrictions
- Configuration example

10.1 EAPS Introduction

EAPS stands for Ethernet Automatic Protection Switching. EAPS uses standard Ethernet and VLAN technology to provide a loop topology and a loop recovery mechanism. If the ring fails, EAPS can restore data communication within 1 second. The number of nodes running EAPS is not limited, and the EAPS recovery time is not limited. EAPS does not depend on other devices; that is, the EAPS ring can contain devices that do not support the EAPS protocol.

10.2 EAPS Basic Concepts

The following are some of the basic concepts used in EAPS:

1. EAPS domain: in a network, an EAPS domain runs on a separate ring. It is a series of node devices that form a separate loop. An EAPS domain includes a master node and one or more transit nodes.
2. A master node is a switch running EAPS, or an EAPS node device. An EAPS domain has only one master node.
3. A transit node is a switch running EAPS, or an EAPS node device, other than the master node in an EAPS domain.
4. Primary port: the port that connects EAPS node devices in an EAPS domain. In an EAPS domain, a node device has one and only one primary port connected to the ring.
5. Secondary port: a port that connects EAPS node devices in an EAPS domain. In an EAPS domain, a node device has only one secondary port connected to the ring.
6. Control VLAN: the VLAN responsible for transmitting the protocol packets of an EAPS domain. There is only one control VLAN in an EAPS domain. 7. Protected VLAN: the VLAN that transmits service data in an EAPS domain. An EAPS domain must have one or more protected VLANs.

10.3 EAPS Protocol Introduction

An EAPS domain runs on an EAPS ring. An EAPS domain contains a master node and one or more transit nodes. Each EAPS node contains the same control VLAN and multiple protected VLANs. In an EAPS domain, each EAPS node contains a primary port and a secondary port. These two ports belong to the control VLAN and to all protected VLANs of the ring. Connect all nodes in the EAPS domain through the primary port and secondary port of each EAPS node device to form an EAPS ring.

Under normal circumstances, when all primary ports and secondary ports in the EAPS domain are link up, the secondary port of the master node is blocked (the secondary port state is set to blocking), which eliminates the loop for service data in the EAPS domain. When the EAPS domain fails, the secondary port of the master node is opened immediately (the secondary port state is set to forwarding) so that it forwards service data and normal service data forwarding is restored.

A transit node handles its primary port and secondary port in the same way

The following describes two kinds of fault checking and loop recovery in EAPS:

10.3.1 Link-Down Alarm

When a transit node finds that its primary port or secondary port is link down, it immediately sends a link-down protocol packet in the control VLAN to the master node through the other link-up port.

When the master node receives the link-down protocol packet:

The master node immediately moves from the complete state to the failed state, opens the secondary port (sets the secondary port state to forwarding), refreshes its Layer 2 and Layer 3 forwarding, and sends a ring-down-flush-fdb packet to notify the other transit nodes in the EAPS domain to refresh their forwarding tables and relearn Layer 2 and Layer 3 forwarding.

When the master node finds that its local primary port is link down, it performs the same operation as when it receives the link-down protocol packet.

When the master node finds that its local secondary port is link down, the master node immediately moves from the complete state to the failed state, refreshes its Layer 2 and Layer 3 forwarding, and sends the ring-down-flush-fdb protocol packet to notify the other transit nodes in the EAPS domain to refresh their forwarding tables and relearn Layer 2 and Layer 3 forwarding.

10.3.2 Loop Check

The master node sends a health protocol packet from its primary port at regular intervals. If the ring is complete, the master node receives the health protocol packet on its secondary port. The master node then restarts its fail period timer, and the state of the master node is complete.

If no health protocol packet is received before the fail period expires, the master node leaves the complete state and enters the failed state, opens the secondary port (sets the secondary port state to forwarding), refreshes its Layer 2 and Layer 3 forwarding, and sends ring-down-flush-fdb to notify the other transit nodes in the EAPS domain to refresh their forwarding tables and relearn Layer 2 and Layer 3 forwarding.

10.3.3 Ring Recovery

The master node sends the health protocol packet from its primary port whether the ring is complete, failed, or in any other state. When the master node is in the failed state and receives a health protocol packet on its secondary port, the ring returns to the complete state. The master node then sets the secondary port state to blocking, refreshes its Layer 2 and Layer 3 forwarding, and sends a ring-up-flush-fdb packet to notify other devices to refresh their Layer 2 and Layer 3 forwarding and relearn Layer 2 and Layer 3 forwarding.

When a port of a transit node returns from link down to link up and the master node finds that the ring is restored, the secondary port of the master node may still be in the forwarding state. This causes a temporary loop. Therefore, when one port of the transit node is in the link up state and the other port changes from link down to link up, the transit node enters a "pre forwarding state". In this state, the port that has just come up is also in the pre forwarding state, so it cannot forward service data, and the possible data loop is broken. When the master node recovers and sends the ring-up-flush-fdb, the transit node receives the protocol packet, switches the node state to the link-up state, sets the port in the pre forwarding state to the forwarding state, and resumes normal forwarding of service data.

If the transit node does not receive the ring-up-flush-fdb protocol packet, it sets the port in the pre forwarding state to the forwarding state after twice the fail time.

10.3.4 EAPS Compatibility With Extreme

The products of Extreme were the first to support EAPS, and the EAPS protocol supported by the switch follows the rfc3619 standard. There are some differences between the EAPS protocol packets of Extreme devices and the packet definitions in rfc3619. The EAPS protocol supported by the switch is fully compatible with Extreme devices, and the compatibility switch is on by default.

10.3.5 Multiple EAPS Domains

The switch supports multiple EAPS domains, 16 in total.

10.4 EAPS Configuration

The basic configuration of the EAPS protocol includes the following basic elements: control VLAN, node mode, primary port, secondary port, protected VLAN, hello time, and fail time. Hello time and fail time are configured by default. Hello time is 1 second and fail timer is 3 seconds.

10.5 Restrictions

1. The primary port must belong to the control VLAN of an EAPS domain and must be a trunk mode member of all protected VLANs.
2. The EAPS protocol cannot run at the same time as the MSTP protocol. If MSTP is started or an MSTP instance is configured, the EAPS protocol cannot be started.
3. When a VLAN starts the vllp protocol, it cannot be configured as the control VLAN or a protected VLAN of EAPS

4. The control VLAN of EAPS can contain only the primary port and the secondary port, and can only be in VLAN trunk mode
5. If a VLAN is configured as the control VLAN of an EAPS domain and that domain has been started, the VLAN cannot be deleted, and its port members cannot be modified or deleted. A control VLAN cannot be configured with Layer 3 interfaces.
6. In a protected VLAN, the primary port and secondary port can only be in trunk mode. Other member ports are not limited.
7. One port can be configured as the primary port or secondary port of only one EAPS domain.
8. The same VLAN can belong to the control VLAN or protected VLAN of only one EAPS domain.
9. The control VLAN of all nodes in an EAPS domain must be the same.

10.6 Brief Introduction to EAPS Commands

To create an EAPS domain, first make sure that the VLAN and port configurations meet the above conditions.

Configuring EAPS requires a certain order. First create an EAPS domain. Before starting the EAPS domain, configure the other parameters according to the requirements above; otherwise the startup does not succeed. To change hello time to a value greater than the current fail time, first change the fail time to a larger number; otherwise the configuration does not succeed. There are no special requirements for the order of the other configurations.

After an EAPS domain has been started, the control VLAN, mode, primary port, and secondary port cannot be modified; the protected VLAN, fail timer, hello time, and Extreme interoperability can be modified.

The primary port and secondary port support LACP ports (that is, trunk groups).

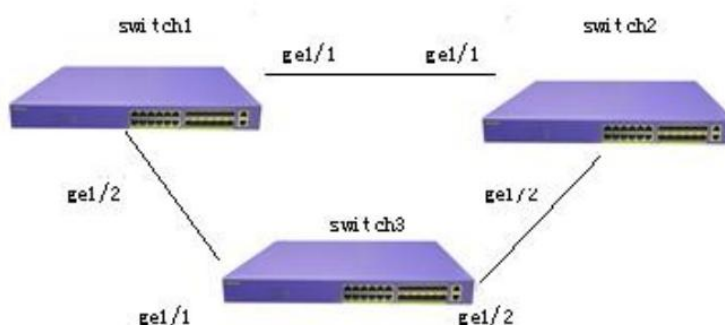
10.6.1 EAPS Configuration Commands

Command	Description	Mode
<code>eaps create <ring-id></code>	Create an EAPS domain	Global Configuration Mode
<code>eaps control-vlan <ring-id> <vlan-id></code>	Configure the control VLAN of an EAPS domain.	Global Configuration Mode
<code>eaps protected-vlan <ring-id> <vlan-id></code>	Add a protected VLAN to an EAPS domain.	Global Configuration Mode
<code>eaps mode <ring-id> <master transit></code>	Configure the node mode in which an EAPS domain runs.	Global Configuration Mode
<code>eaps primary-port <ring-id> <ifname></code>	Configure the primary port of an EAPS domain.	Global Configuration Mode
<code>eaps secondary-port <ring-id> <ifname></code>	Configure the secondary port of an EAPS domain.	Global Configuration Mode

Command	Description	Mode
<code>eaps data-span <ring-id></code>	Configure cross-ring forwarding of EAPS ring data	Global Configuration Mode
<code>eaps fail-time <ring-id> <secs></code>	Configure the timeout of the fail period timer of an EAPS domain. The default is 3 seconds. The unit is seconds.	Global Configuration Mode
<code>eaps hello-time <ring-id> <secs></code>	Configure how often an EAPS domain sends health packets. The default is 1 second. The unit is seconds. The hello timer must be less than the fail time.	Global Configuration Mode
<code>eaps extreme-interoperability <ring-id> <enable disable></code>	Enable or disable compatibility with Extreme devices. Compatibility is enabled by default.	Global Configuration Mode
<code>eaps enable <ring-id></code>	Start an EAPS domain	Global Configuration Mode
<code>eaps disable <ring-id></code>	Shut down an EAPS domain	Global Configuration Mode
<code>show eaps</code>	Display information about the EAPS domains started in the system	Normal Mode / Privileged Mode
<code>Show eaps <ring-id></code>	Display the details of an EAPS domain	Normal Mode / Privileged Mode

10.7 Single Ring Configuration Example

There are three switches: Switch1, Switch2, and Switch3. VLAN 1 does not form a loop when traffic is forwarded through the EAPS protocol. At the same time, the standby link is enabled when one link between Switch1, Switch2, and Switch3 is disconnected. To meet these requirements, configure Switch1 in master mode, and configure Switch2 and Switch3 in transit mode. Add a control VLAN, VLAN 2, for protocol packet transmission.



Switch1 configuration:

Switch1 is configured as the master of EAPS domain ring 1. The control

VLAN is VLAN 2, the protected VLAN is VLAN 1, the primary port is GE1/1, and the secondary port is GE1/2. Other configurations use the default values. Switch#configure terminal

#Add VLANs 2 and 3

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-3
Switch(config-vlan)#exit
```

#Configure GE1/1 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode trunk
Switch(config-ge1/1)#switchport trunk native vlan 3
Switch(config-ge1/1)#switchport trunk allowed vlan add 1,2
```

#Configure GE1/2 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport mode trunk
Switch(config-ge1/2)#switchport trunk native vlan 3
Switch(config-ge1/2)#switchport trunk allowed vlan add 1,2
Switch(config-ge1/2)#exit
Switch(config)#exit
Switch#show vlan
```

VLAN Name State Member ports ([u]-Untagged, [t]-Tagged) -----

```
1 vlan1 active [t]ge1/1 [t]ge1/2 [u]ge1/3 [u]ge1/4 [u]ge1/5 [u]ge1/6 [u]ge1/7 [u]ge1/8 [u]ge1/9
[u]ge1/10 [u]ge1/11 [u]ge1/12
2 vlan2 active [t]ge1/1 [t]ge1/2
3 vlan3 active [u]ge1/1 [u]ge1/2
```

```
Switch#configure terminal
```

Create EAPS domain ring 1

```
Switch(config)#eaps create 1
```

#Configure VLAN 2 as the control VLAN

```
Switch(config)#eaps control-vlan 1 2
```

#Configure VLAN 1 as a protected VLAN

```
Switch(config)#eaps protected-vlan 1 1
```

#Configure Switch1 as the master node

```
Switch(config)#eaps mode 1 master
```

#Configure GE1/1 as the primary port

```
Switch(config)#eaps primary-port 1 ge1/1
```

#Configure GE1/2 as the secondary port

```
Switch(config)#eaps secondary-port 1 ge1/2
```

#Start EAPS domain ring 1

```
Switch(config)#eaps enable 1
```

Switch2 configuration

Switch2 is configured as the transit node of EAPS domain ring 1. The control VLAN is VLAN 2, the protected VLAN is VLAN 1, the primary port is GE1/1, and the secondary port is GE1/2. Other configurations use the default values.

```
Switch#configure terminal
```

#Add VLANs 2 and 3

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-3
Switch(config-vlan)#exit
```

#Configure GE1/1 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode trunk
Switch(config-ge1/1)#switchport trunk native vlan 3
Switch(config-ge1/1)#switchport trunk allowed vlan add 1,2
```

#Configure GE1/2 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport mode trunk
Switch(config-ge1/2)#switchport trunk native vlan 3
Switch(config-ge1/2)#switchport trunk allowed vlan add 1,2
Switch(config-ge1/2)#exit
Switch(config)#exit
Switch#show vlan
```

VLAN Name State Member ports ([u]-Untagged, [t]-Tagged) ---- -----

```
1 vlan1 active [t]ge1/1 [t]ge1/2 [u]ge1/3 [u]ge1/4 [u]ge1/5 [u]ge1/6 [u]ge1/7 [u]ge1/8 [u]ge1/9
[u]ge1/10
2 vlan2 active [t]ge1/1 [t]ge1/2
3 vlan3 active [u]ge1/1 [u]ge1/2
```

```
Switch#configure terminal
```

#Create EAPS domain ring 1

```
Switch(config)#eaps create
```

#Configure VLAN 2 as the control VLAN

```
Switch(config)#eaps control-vlan 1 2
```

#Configure VLAN 1 as a protected VLAN

```
Switch(config)#eaps protected-vlan 1
```

#Configure the switch as a transit node

```
Switch(config)#eaps mode 1 transit
```

#Configure GE1/1 as the primary port

```
Switch(config)#eaps primary-port 1 ge1/1
```

#Configure GE1/2 as the secondary port


```
Switch(config)#eaps secondary-port 1 ge1/2
```

#Start EAPS domain ring 1

```
Switch(config)#eaps enable 1
```

Switch3 configuration

Switch3 is configured as the transit node of EAPS domain ring 1. The control VLAN is VLAN 2, the protected VLAN is VLAN 1, the primary port is GE1/1, and the secondary port is GE1/2. Other configurations use the default values.

```
Switch#configure terminal
```

Add VLANs 2 and 3

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-3
Switch(config-vlan)#exit
```

Configure GE1/1 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode trunk
Switch(config-ge1/1)#switchport trunk native vlan 3
Switch(config-ge1/1)#switchport trunk allowed vlan add 1,2
```

Configure GE1/2 as a trunk member of VLAN 1 and VLAN 2.

```
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport mode trunk
Switch(config-ge1/2)#switchport trunk native vlan 3
Switch(config-ge1/2)#switchport trunk allowed vlan add 1,2
Switch(config-ge1/2)#exit
Switch(config)#exit
Switch#show vlan
```

VLAN Name State Member ports ([u]-Untagged, [t]-Tagged) ----

1 vlan1 active [t]ge1/1 [t]ge1/2 [u]ge1/3 [u]ge1/4 [u]ge1/5 [u]ge1/6 [u]ge1/7 [u]ge1/8 [u]ge1/9

[u]ge1/10 [u]ge1/11 [u]ge1/12

2 vlan2 active [t]ge1/1 [t]ge1/2

3 vlan3 active [u]ge1/1 [u]ge1/2

```
Switch#configure terminal
```

#Create EAPS domain ring 1

```
Switch(config)#eaps create 1
```

#Configure VLAN 2 as the control VLAN

```
Switch(config)#eaps control-vlan 1
```

#Configure VLAN 1 as a protected VLAN

```
Switch(config)#eaps protected-vlan 1 1
```

#Configure Switch3 as a transit node

```
Switch(config)#eaps mode 1 transit
```

#Configure GE1/1 as the primary port Switch(config)#eaps primary-port 1 ge1/1

#Configure GE1/2 as the secondary port

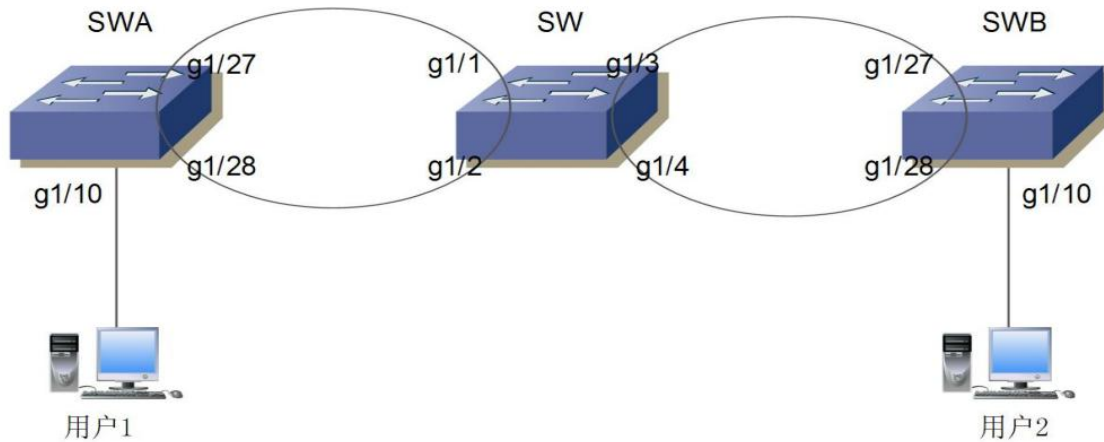
```
Switch(config)#eaps secondary-port 1 ge1/2
```

#Start EAPS domain ring 1

```
Switch(config)#eaps enable 1
```

10.8 Cross Ring Data Forwarding Configuration Example

There are three switches, SWA, SW, and SWB, which use the EAPS protocol to let vlan1 and vlan2 communicate across the rings. The topology is as follows:



On SWA, ring 1 controls VLAN 111 and protects VLAN 1 and VLAN 2. The configuration is as follows:

```
vlan database
vlan 2
vlan 111
interface ge1/10
switchport access vlan 2
interface ge1/27
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 111

```
interface ge1/28
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 111

```
eaps create 1
eaps mode 1 Transit
eaps primary-port 1 ge1/27
eaps secondary-port 1 ge1/28
eaps control-vlan 1 111
eaps protected-vlan 1 1
eaps protected-vlan 1 2
eaps enable 1
```

On SW, ring 1 connects to SWA, controls VLAN 111, and protects VLAN 1 and VLAN 2. Ring 2 connects to SWB, controls VLAN 222, and protects VLAN 3333 (a virtual VLAN, to which the interface must be added). To forward the data of ring 1 and ring 2 across the rings, configure the EAPS data span command. The configuration is as follows:

```
vlan database
vlan 2
vlan 111
vlan 222
vlan 3333
interface ge1/1
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 111

```
interface ge1/2
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 111

```
interface ge1/3
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 222 switchport trunk allowed vlan add 3333 ###Add virtual vlan 3333

```
interface ge1/4
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 222 switchport trunk allowed vlan add 3333 ###Add virtual vlan 3333

```
eaps create 1
eaps mode 1 Master
eaps primary-port 1 ge1/1
eaps secondary-port 1 ge1/2
eaps control-vlan 1 111
eaps protected-vlan 1 1
eaps protected-vlan 1 2
eaps data-span 1
eaps enable 1
eaps create 2
eaps mode 2 Transit
eaps primary-port 2 ge1/3
eaps secondary-port 2 ge1/4
eaps control-vlan 2 222
```

eaps protected-vlan 2 3333 ###This is the virtual protected vlan

```
eaps data-span 2
eaps enable 2
```

On SWB, ring 2 connects to ring 2 of SW, controls VLAN 222, and protects VLAN 1 and VLAN 2. The configuration is as follows:

```
vlan database
vlan 2
vlan 222
interface ge1/10
switchport access vlan 2
interface ge1/27
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 222

```
interface ge1/28
switchport mode trunk
```

switchport trunk allowed vlan add 2 switchport trunk allowed vlan add 222

```
eaps create 2
eaps mode 2 Master
eaps primary-port 2 ge1/27
eaps secondary-port 2 ge1/28
eaps control-vlan 2 222
eaps protected-vlan 2 1
eaps protected-vlan 2 2
eaps enable 2
```

After this configuration, vlan1 and vlan2 can communicate with each other. The EAPS node mode can be modified as required.

Chapter 11 Configuring ERPS

11.1 ERPS Overview

ERPS (Ethernet Ring Protection Switching) is a ring protection protocol developed by the ITU, also known as G.8032. It is a link layer protocol designed specifically for Ethernet ring networks. It prevents the broadcast storm caused by a data loop when the Ethernet ring network is complete, and it quickly restores communication between the nodes on the ring network when one link on the Ethernet ring network is broken. The ERPS protocol provides a fast protection mechanism for Ethernet ring networks. It quickly restores network transmission when the ring network fails, which ensures the high availability and reliability of the switch in a ring network topology.

11.2 Introduction to ERPS Technology

11.2.1 ERPS Ring

The ERPS ring is based on the principle of minimizing the ring: each ring must be the smallest ring. Rings are divided into a main ring and a sub-ring. The main ring is a closed ring; the sub-ring is an unclosed ring or a closed ring. All of them must be configured by command.

Each ERPS ring, whether a main ring or a sub-ring, has five states: (1) Idle state: the state when each physical link in the ring network is connected; (2) Protection state: the state when one or more physical links in the ring network are disconnected; (3) Manual switch state: the ring state is changed manually; (4) Forced switch state: the ring state is changed by force; (5) Pending state: a pending intermediate state.

11.2.2 ERPS Node

The Layer 2 switching devices that join an ERPS ring are called nodes. No more than two ports of each node can join the same ERPS ring: one port is an RPL port and the other is a normal ring port.

Globally, node roles fall into the following two types: (1) Intersecting nodes: in an intersecting ERPS ring, nodes that belong to multiple rings at the same time are called intersecting nodes; (2) Non-intersecting nodes: in an intersecting ERPS ring, nodes that belong to only one ERPS ring are called non-intersecting nodes.

The node modes specified in the ERPS protocol are the RPL owner node, the RPL neighbor node, and the ordinary ring node.

(1) RPL owner node: there is only one RPL owner node in an ERPS ring, and it is determined by user configuration. Its RPL port is blocked to prevent loops in the ERPS ring. When the RPL owner node receives a failure message and learns that another node or link on the ERPS ring has failed, it automatically releases the RPL port. That port resumes receiving and sending traffic so that traffic is not interrupted;

(2) RPL neighbor node: a node that is directly connected to the RPL port of the RPL owner node. Normally, both the RPL port of the RPL owner node and the RPL port of the RPL neighbor node are blocked to prevent loops. When the ERPS ring fails, the RPL port of the RPL owner node and the RPL port of the RPL neighbor node are both released;

(3) Ordinary ring nodes: in the ERPS ring, all nodes except the RPL owner node and the RPL neighbor node are ordinary ring nodes. On an ordinary ring node, the RPL port is no different from an ordinary ring port. The ring ports of an ordinary ring node monitor the state of the ERPS protocol links directly connected to the node, and they inform the other nodes of link state changes in time;

11.2.3 Link and Channel

(1) RPL (Ring Protection Link): each ERPS ring has only one RPL, that is, the link on which the RPL port of the RPL owner node is located. When the Ethernet ring is in the Idle state, the RPL link is in the blocking state and does not forward data packets, which avoids forming a loop;

(2) Subring link: a link that belongs to a subring and is controlled by the subring in an intersecting ring;

(3) RAPS (Ring Auto Protection Switch) virtual channel: in an intersecting ring, the path between intersecting nodes that transmits sub-ring protocol messages but does not belong to the sub-ring is called the RAPS virtual channel of the sub-ring.

11.2.4 ERPS VLAN

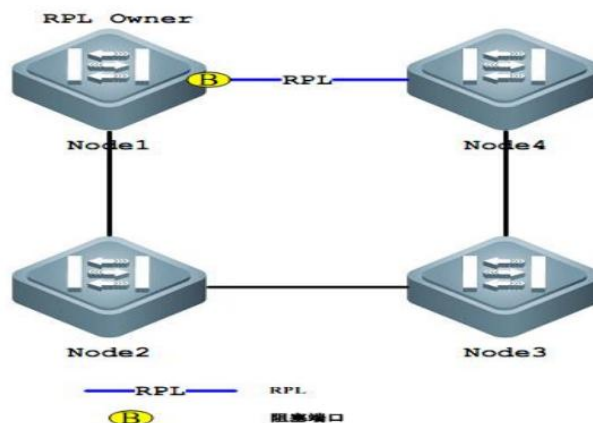
There are two types of VLANs in ERPS:

(1) RAPS VLAN: it is used to transmit ERPS protocol messages. All ports on the device that access the ERPS ring belong to the RAPS VLAN, and only ports that access the ERP ring can join this VLAN. The RAPS VLAN must be different for different rings. You cannot configure an IP address on the interface of a RAPS VLAN;

(2) Data VLAN: unlike the RAPS VLAN, the data VLAN is used to transmit data packets. A data VLAN can include both ERP ring ports and non-ERP ring ports.

11.3 ERPS Operating Principle

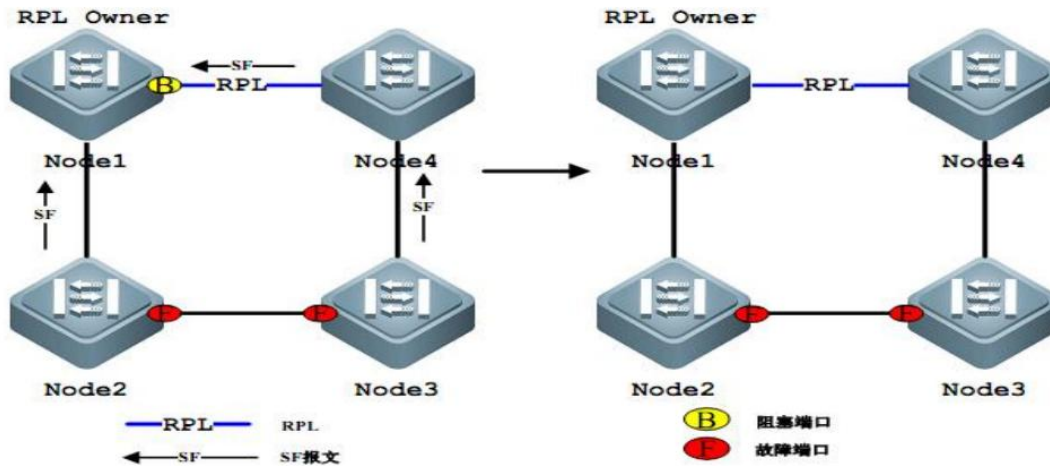
11.3.1 Normal State



(1) All nodes are connected in a ring in the physical topology;

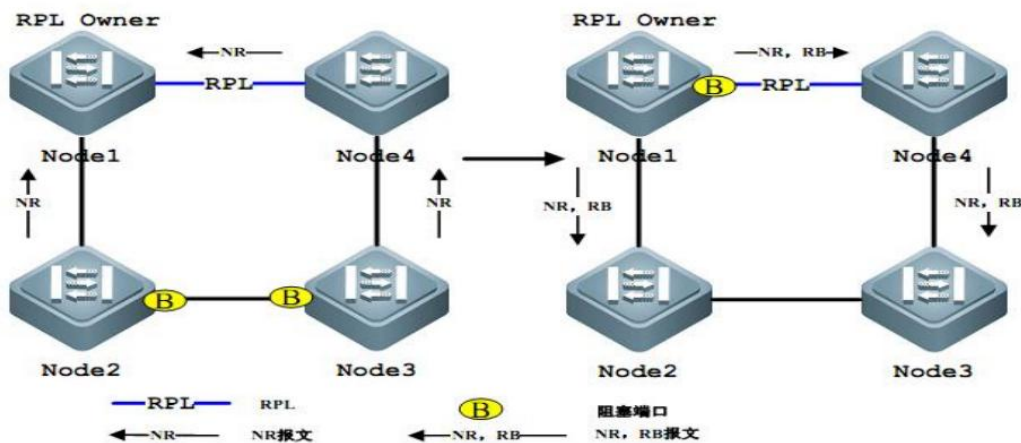
- (2) The loop protection protocol blocks the RPL link so that no loop is formed. As shown in the figure above, the link between Node1 and Node4 is an RPL link;
- (3) Fault detection is performed on each link between adjacent nodes.

11.3.2 Link Failure



- (1) The node adjacent to the failed link blocks the failed link and reports the failure to the other nodes on the ring with a RAPS (SF) message, as shown in the figure above. Suppose the link between Node2 and Node3 fails. Node2 and Node3 block the failed link after the holdoff timer expires, and each sends a RAPS (SF) message to every node on the ring network;
- (2) The RAPS (SF) message triggers the RPL owner node to open the RPL port. The RAPS (SF) message also triggers all nodes to update their own MAC entries, and the nodes then enter the protection state.

11.3.3 Link Recovery

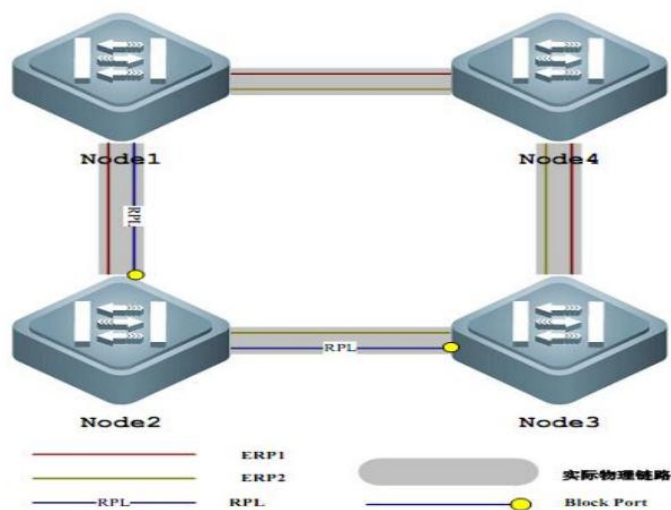


- (1) When the fault is recovered, the node adjacent to the fault stays in the blocking state and sends a RAPS (NR) message indicating that there is no local fault;

- (2) After the guard timer expires, the RPL Owner node starts the WTR timer when it receives the first RAPS (NR) message;
- (3) When the WTR timer expires, the RPL Owner node blocks the RPL and sends a RAPS (NR, RB) message;
- (4) After receiving that message, the other nodes update their own MAC table entries, and the node that sent the RAPS (NR) message stops sending the message periodically and opens the port that was blocked. The ring network has returned to its original normal state.

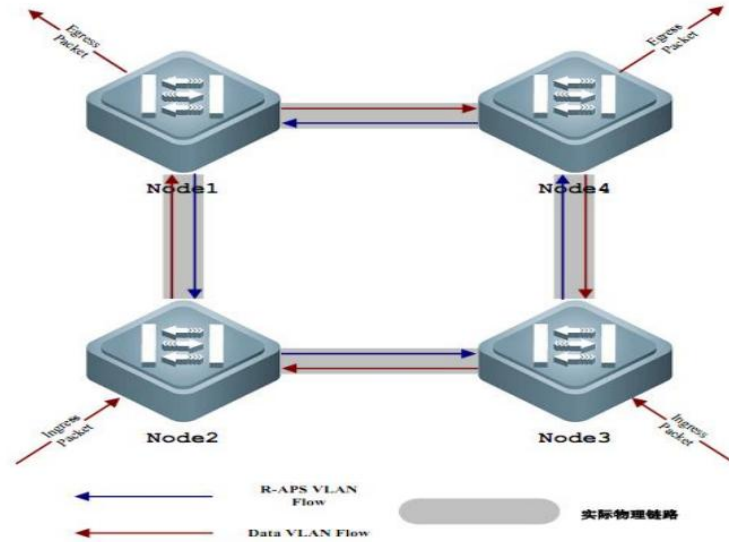
11.4 Technical Features of ERPS

11.4.1 ERPS Load Balancing



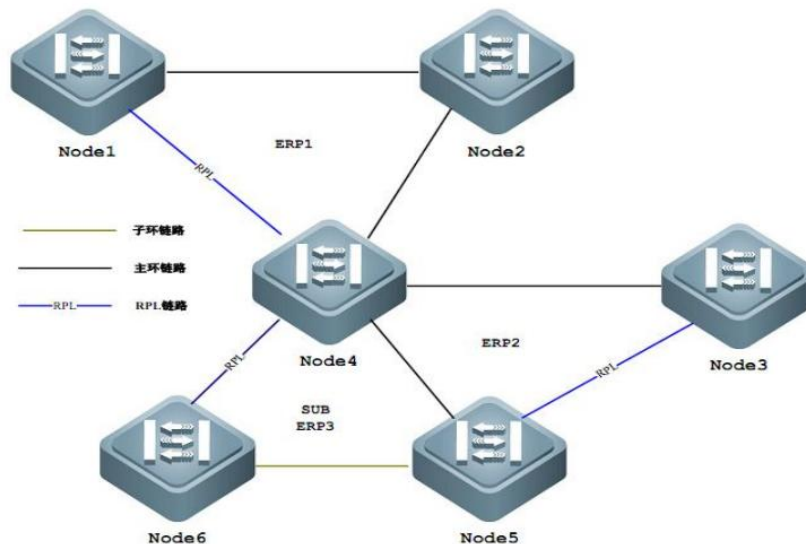
Multiple instances and multiple ERPS rings are configured on the same physical ring network, and different ERPS rings carry the traffic of different VLANs (called protection VLANs), so that the data traffic of different VLANs follows different topologies in the ring network and load sharing is achieved. As shown in the figure above, one physical ring network corresponds to two instances and two ERPS rings. The VLANs protected by the two ERPS rings are different. Node2 is the RPL owner node of ERP 1, and Node3 is the RPL owner node of ERP 2. Through configuration, different VLANs can block different links, which achieves load sharing on a single ring.

11.4.2 Good Security



There are two types of VLANs in ERPS: the RAPS VLAN and the data VLAN. The RAPS VLAN is used only to transmit ERPS protocol messages, and ERPS processes protocol messages only from the RAPS VLAN. It does not process any protocol attack message from the data VLAN, which improves the security of ERPS.

11.4.3 Support for Multi-Ring Intersection and Tangency



190

As shown in the figure above, ERPS supports joining multiple rings at the same node (Node4) in a tangent or intersecting form, which greatly increases networking flexibility.

11.5 ERPS Protocol Commands

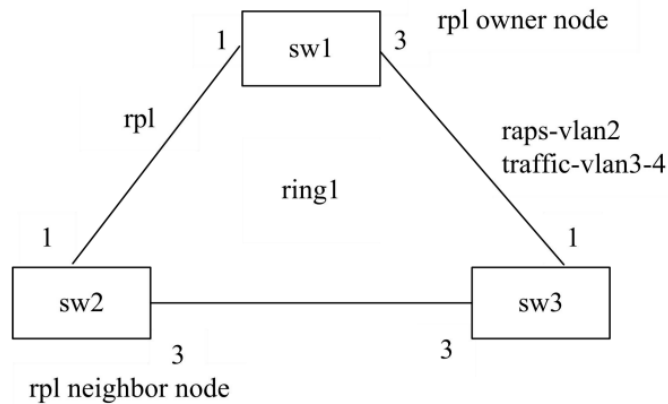
Command	Description	CLI Mode
<code>erps predefine configuration (ring-node rpl-owner-node)</code>	Enable ERPS predefined configuration	Global Configuration Mode
<code>no erps predefine configuration</code>	Disable ERPS predefined configuration	Global Configuration Mode
<code>erps <1-8></code>	Create an ERPS instance	Global Configuration Mode
<code>no erps <1-8></code>	Delete an ERPS instance	Global Configuration Mode
<code>node-role (interconnection none-interconnection)</code>	Configure the role of the node in the ERPS ring: interlink or non-interlink	ERPS Mode
<code>ring <1-32></code>	Create an ERPS ring	ERPS Mode
<code>no ring <1-32></code>	Delete an ERPS ring	ERPS Mode
<code>ring <1-32> ring-mode (major-ring sub-ring)</code>	Configure the ERPS ring mode: primary or subring	ERPS Mode
<code>ring <1-32> node-mode (rpl-owner-node rpl-neighbor-node ring-node)</code>	Configure the ERPS ring node mode: RPL owner node, RPL neighbor node, or normal ring node	ERPS Mode
<code>ring <1-32> raps-vlan <2-4094></code>	Configure the ERPS ring protocol VLAN	ERPS Mode
<code>no ring <1-32> raps-vlan</code>	Delete an ERPS ring protocol VLAN	ERPS Mode
<code>ring <1-32> traffic-vlan <1-4094></code>	Configure the ERPS ring data VLAN	ERPS Mode
<code>no ring <1-32> traffic-vlan <1-4094></code>	Delete the ERPS ring data VLAN	ERPS Mode
<code>ring <1-32> (rpl-port r1-port) IFNAME</code>	Configure the ERPS ring port: RPL port or normal ring port	ERPS Mode
<code>no ring <1-32> (rpl-port r1-port)</code>	Configure the ERPS ring port: RPL port or normal ring port	ERPS Mode
<code>ring <1-32> revertive-behaviour (revertive non-revertive)</code>	Configure the ERPS ring recovery behavior: recoverable or non-recoverable	ERPS Mode
<code>ring <1-32> hold-off-time <0-10000></code>	Configure the ERPS ring hold-off time	ERPS Mode
<code>no ring <1-32> hold-off-time</code>	Restore the default ERPS ring hold-off time	ERPS Mode
<code>ring <1-32> guard-time <10-2000></code>	Configure the ERPS ring guard time	ERPS Mode

Command	Description	CLI Mode
no ring <1-32> guard-time	Restore the default time for the ERPS ring	ERPS Mode
ring <1-32> wtr-time <1-12>	Configure the ERPS ring wtr time	ERPS Mode
no ring <1-32> wtr-time	Restore the default ERPS ring wtr time	ERPS Mode
ring <1-32> wtb-time <1-10>	Configure the ERPS ring WTB time	ERPS Mode
no ring <1-32> wtb-time	Restore the default ERPS ring WTB time	ERPS Mode
ring <1-32> raps-send-time <1-10>	Configure the sending time of ERPS ring protocol messages	ERPS Mode
no ring <1-32> raps-send-time	Restore the default sending time of ERPS ring protocol messages	ERPS Mode
ring <1-32> (enable disable)	Turn the ERPS ring on or off	ERPS Mode
ring <1-32> forced-switch IFNAME	Force a switch on an ERPS ring port	ERPS Mode
ring <1-32> clear forced-switch	Clear the forced switch of an ERPS ring	ERPS Mode
ring <1-32> manual-switch IFNAME	Manually switch an ERPS ring port	ERPS Mode
ring <1-32> clear manual-switch	Clear the manual switch of an ERPS ring	ERPS Mode
ring <1-32> clear recovery	Manual recovery when the non-recoverable behavior of the ERPS ring is cleared, or manual recovery before WTR/WTB expires	ERPS Mode
show erps	Display a summary of all ERPS instances and rings on the device	Privileged Mode
show erps <1-8>	Display the details of a single ERPS instance and ring on the device	Privileged Mode

11.6 ERPS Typical Application

11.6.1 Single Ring Example

As shown in the figure below, nodes sw1, sw2, and sw3 form ERPS single ring ring1. Ports 1 and 3 of each node are used as ERPS ring ports. The protocol VLAN of the ring is 2, and the data VLANs are 3 and 4. The sw1 node is the RPL owner node, and the sw2 node is the RPL neighbor node. The link between sw1 and sw2 is an RPL link.



(1) Configure SW 1:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-owner-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

(2) Configure SW 2:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-neighbor-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

(3) Configure SW 3:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```

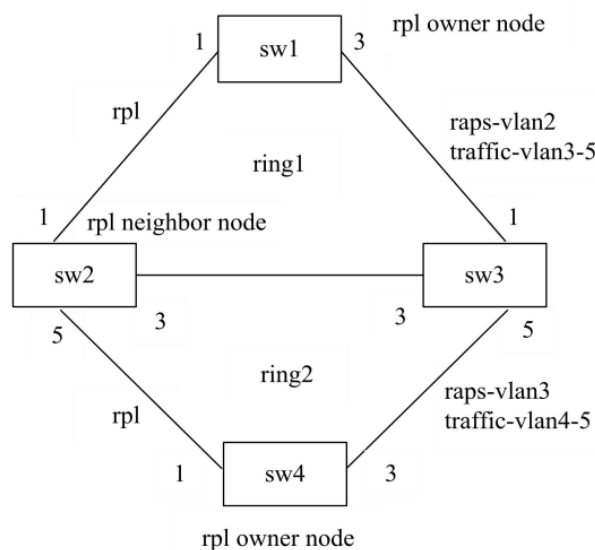
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode ring-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit

```

11.6.2 Multi-Ring Example

As shown in the figure below, nodes sw1, sw2, and sw3 form ERPS main ring ring1. Ports 1 and 3 of nodes sw1, sw2, and sw3 serve as the ring ports of main ring ring1. The protocol VLAN of main ring ring1 is 2, and the data VLANs are 3, 4, and 5. The sw1 node is the RPL owner node of main ring ring1, the sw2 node is the RPL neighbor node of main ring ring1, and the link between sw1 and sw2 is the RPL link of main ring ring1.

Nodes sw2, sw3, and sw4 form ERPS subring ring2. Port 5 of nodes sw2 and sw3 and ports 1 and 3 of node sw4 are used as the ring ports of subring ring2. The protocol VLAN of subring ring2 is 3, and the data VLANs are 4 and 5. The sw4 node is the RPL owner node of subring ring2, and the link between sw2 and sw4 is the RPL link of subring ring2.



(1) Configure SW 1:

```

Switch>enable
Switch#configure terminal

```


Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-5
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS primary ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-owner-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 traffic-vlan 5
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

(2) Configure SW 2:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-5
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)#exit
Switch(config)# interface ge1/5
Switch(config-ge1/5)# switchport mode trunk
Switch(config-ge1/5)# switchport trunk allowed vlan add 3
Switch(config-ge1/5)# switchport trunk allowed vlan add 4
Switch(config-ge1/5)# switchport trunk allowed vlan add 5
Switch(config-ge1/5)#exit
```

Configure ERPS instance 1, ERPS primary ring 1, subring 2

```
Switch(config)#erps 1
Switch(config-erps-1)# node-role interconnection
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-neighbor-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 traffic-vlan 5
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#ring 2
Switch(config-erps-1)# ring 2 ring-mode sub-ring
Switch(config-erps-1)# ring 2 node-mode ring-node
Switch(config-erps-1)# ring 2 raps-vlan 3
Switch(config-erps-1)# ring 2 traffic-vlan 4
Switch(config-erps-1)# ring 2 traffic-vlan 5
Switch(config-erps-1)# ring 2 rpl-port ge1/5
Switch(config-erps-1)# ring 2 enable
Switch(config-erps-1)#exit
```

(3) Configure SW 3:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-5
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)#exit
Switch(config)# interface ge1/5
Switch(config-ge1/5)# switchport mode trunk
Switch(config-ge1/5)# switchport trunk allowed vlan add 3
Switch(config-ge1/5)# switchport trunk allowed vlan add 4
Switch(config-ge1/5)# switchport trunk allowed vlan add 5
Switch(config-ge1/5)#exit
```

Configure ERPS instance 1, ERPS primary ring 1, subring 2

```
Switch(config)#erps 1
Switch(config-erps-1)# node-role interconnection
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode ring-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 traffic-vlan 5
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#ring 2
Switch(config-erps-1)# ring 2 ring-mode sub-ring
Switch(config-erps-1)# ring 2 node-mode ring-node
Switch(config-erps-1)# ring 2 raps-vlan 3
Switch(config-erps-1)# ring 2 traffic-vlan 4
Switch(config-erps-1)# ring 2 traffic-vlan 5
Switch(config-erps-1)# ring 2 rpl-port ge1/5
Switch(config-erps-1)# ring 2 enable
Switch(config-erps-1)#exit
```

(4) Configure SW 4:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 3-5
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)#exit
```

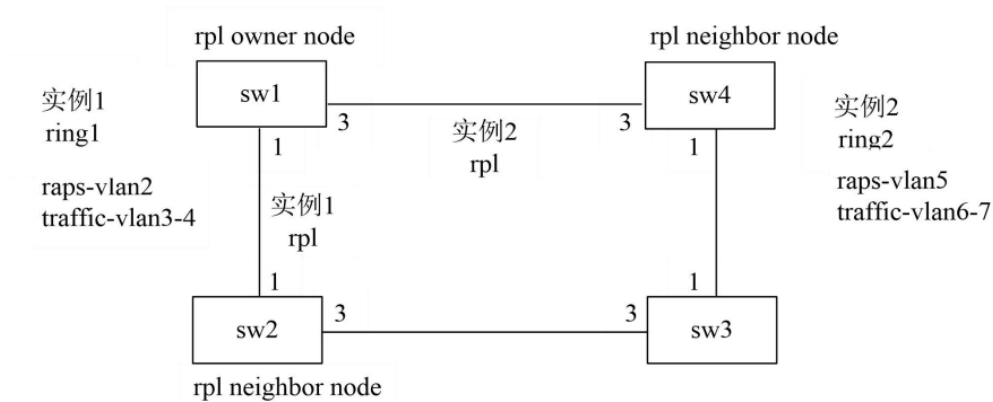
Configure ERPS instance 1, ERPS subring 2

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 2
Switch(config-erps-1)# ring 2 ring-mode sub-ring
Switch(config-erps-1)# ring 2 node-mode rpl-owner-node
Switch(config-erps-1)# ring 2 raps-vlan 3
Switch(config-erps-1)# ring 2 traffic-vlan 1
Switch(config-erps-1)# ring 2 traffic-vlan 4
Switch(config-erps-1)# ring 2 traffic-vlan 5
Switch(config-erps-1)# ring 2 rpl-port ge1/1
Switch(config-erps-1)# ring 2 rl-port ge1/3
Switch(config-erps-1)# ring 2 enable
Switch(config-erps-1)#exit
```

11.6.3 Multi-Instance Load Balancing Example

As shown in the figure below, nodes sw1, sw2, sw3, and sw4 form single ring 1 of ERPS instance 1. Ports 1 and 3 of each node are used as the ERPS ring ports. The protocol VLAN of the ring is 2, and the data VLANs are 3 and 4. Node sw1 is the RPL owner node, and node sw2 is the RPL neighbor node. The link between sw1 and sw2 is an RPL link.

Nodes sw1, sw2, sw3, and sw4 form single ring 2 of ERPS instance 2. Ports 1 and 3 of each node are used as the ERPS ring ports. The protocol VLAN of the ring is 5, and the data VLANs are 6 and 7. Node sw1 is the RPL owner node, and node sw4 is the RPL neighbor node. The link between sw1 and sw4 is an RPL link.



(1) Configuration example 1:

Configure SW 1:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-owner-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

Configure SW 2:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode rpl-neighbor-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

Configure SW 3:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1

```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode ring-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

Configure SW 4:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 2-4
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 2
Switch(config-ge1/1)# switchport trunk allowed vlan add 3
Switch(config-ge1/1)# switchport trunk allowed vlan add 4
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 2
Switch(config-ge1/3)# switchport trunk allowed vlan add 3
Switch(config-ge1/3)# switchport trunk allowed vlan add 4
Switch(config-ge1/3)#exit
```

Configure ERPS instance 1, ERPS single ring 1


```
Switch(config)#erps 1
Switch(config-erps-1)#ring 1
Switch(config-erps-1)# ring 1 ring-mode major-ring
Switch(config-erps-1)# ring 1 node-mode ring-node
Switch(config-erps-1)# ring 1 raps-vlan 2
Switch(config-erps-1)# ring 1 traffic-vlan 1
Switch(config-erps-1)# ring 1 traffic-vlan 3
Switch(config-erps-1)# ring 1 traffic-vlan 4
Switch(config-erps-1)# ring 1 rpl-port ge1/1
Switch(config-erps-1)# ring 1 rl-port ge1/3
Switch(config-erps-1)# ring 1 enable
Switch(config-erps-1)#exit
```

(2) Configuration example 2:

Configure SW 1:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 5-7
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)# switchport trunk allowed vlan add 6
Switch(config-ge1/1)# switchport trunk allowed vlan add 7
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)# switchport trunk allowed vlan add 6
Switch(config-ge1/3)# switchport trunk allowed vlan add 7
Switch(config-ge1/3)#exit
```

Configure ERPS instance 2, ERPS single ring 2

```
Switch(config)#erps 2
Switch(config-erps-2)#ring 2
Switch(config-erps-2)# ring 2 ring-mode major-ring
Switch(config-erps-2)# ring 2 node-mode rpl-owner-node
Switch(config-erps-2)# ring 2 raps-vlan 5
Switch(config-erps-2)# ring 2 traffic-vlan 1
Switch(config-erps-2)# ring 2 traffic-vlan 6
Switch(config-erps-2)# ring 2 traffic-vlan 7
Switch(config-erps-2)# ring 2 rpl-port ge1/3
Switch(config-erps-2)# ring 2 rl-port ge1/1
Switch(config-erps-2)# ring 2 enable
Switch(config-erps-2)#exit
```

Configure SW 2:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 5-7
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)# switchport trunk allowed vlan add 6
Switch(config-ge1/1)# switchport trunk allowed vlan add 7
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)# switchport trunk allowed vlan add 6
Switch(config-ge1/3)# switchport trunk allowed vlan add 7
Switch(config-ge1/3)#exit
```

Configure ERPS instance 2, ERPS single ring 2

```
Switch(config)#erps 2
Switch(config-erps-2)#ring 2
Switch(config-erps-2)# ring 2 ring-mode major-ring
Switch(config-erps-2)# ring 2 node-mode ring-node
Switch(config-erps-2)# ring 2 raps-vlan 5
Switch(config-erps-2)# ring 2 traffic-vlan 1
Switch(config-erps-2)# ring 2 traffic-vlan 6
Switch(config-erps-2)# ring 2 traffic-vlan 7
Switch(config-erps-2)# ring 2 rpl-port ge1/1
Switch(config-erps-2)# ring 2 rl-port ge1/3
Switch(config-erps-2)# ring 2 enable
Switch(config-erps-2)#exit
```

Configure SW 3:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 5-7
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)# switchport trunk allowed vlan add 6
Switch(config-ge1/1)# switchport trunk allowed vlan add 7
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)# switchport trunk allowed vlan add 6
Switch(config-ge1/3)# switchport trunk allowed vlan add 7
Switch(config-ge1/3)#exit
```

Configure ERPS instance 2, ERPS single ring 2

```
Switch(config)#erps 2
Switch(config-erps-2)#ring 2
Switch(config-erps-2)# ring 2 ring-mode major-ring
Switch(config-erps-2)# ring 2 node-mode ring-node
Switch(config-erps-2)# ring 2 raps-vlan 5
Switch(config-erps-2)# ring 2 traffic-vlan 1
Switch(config-erps-2)# ring 2 traffic-vlan 6
Switch(config-erps-2)# ring 2 traffic-vlan 7
Switch(config-erps-2)# ring 2 rpl-port ge1/1
Switch(config-erps-2)# ring 2 rl-port ge1/3
Switch(config-erps-2)# ring 2 enable
Switch(config-erps-2)#exit
```

Configure SW 4:

```
Switch>enable
Switch#configure terminal
```

Create the ERPS protocol and data VLANs

```
Switch(config)#vlan database
Switch(config-vlan)#vlan 5-7
Switch(config-vlan)#exit
```

Configure the VLAN mode of the ring port as trunk, and add the ERPS protocol and data VLANs

```
Switch(config)# interface ge1/1
Switch(config-ge1/1)# switchport mode trunk
Switch(config-ge1/1)# switchport trunk allowed vlan add 5
Switch(config-ge1/1)# switchport trunk allowed vlan add 6
Switch(config-ge1/1)# switchport trunk allowed vlan add 7
Switch(config-ge1/1)#exit
Switch(config)# interface ge1/3
Switch(config-ge1/3)# switchport mode trunk
Switch(config-ge1/3)# switchport trunk allowed vlan add 5
Switch(config-ge1/3)# switchport trunk allowed vlan add 6
Switch(config-ge1/3)# switchport trunk allowed vlan add 7
Switch(config-ge1/3)#exit
```

Configure ERPS instance 2, ERPS single ring 2

```
Switch(config)#erps 2
Switch(config-erps-2)#ring 2
Switch(config-erps-2)# ring 2 ring-mode major-ring
Switch(config-erps-2)# ring 2 node-mode rpl-neighbor-node
Switch(config-erps-2)# ring 2 raps-vlan 5
Switch(config-erps-2)# ring 2 traffic-vlan 1
Switch(config-erps-2)# ring 2 traffic-vlan 6
Switch(config-erps-2)# ring 2 traffic-vlan 7
Switch(config-erps-2)# ring 2 rpl-port ge1/3
Switch(config-erps-2)# ring 2 rl-port ge1/1
Switch(config-erps-2)# ring 2 enable
Switch(config-erps-2)#exit
```

Chapter 12 Configuring AAA

This chapter describes how to configure 802.1X and RADIUS on the switch to prevent unauthorized users from accessing the network. For the use of the 802.1X client and HyperBoss, refer to their respective instruction manuals. This chapter includes the following contents:

- 802.1x Introduction
- RADIUS Introduction
- Configuring 802.1x
- Configuring RADIUS

AAA stands for Authentication, Authorization, and Accounting. It provides a consistent framework for configuring the authentication, authorization, and accounting security functions. Configuring AAA is a form of network security management, and network security here mainly refers to access control: which users can access the network, what services are available to users with access, and how users of network resources are billed.

Authentication: verifies whether the user can gain access.

Authorization: what services are available to authorized users.

Accounting: records the use of network resources by users.

Network companies have launched a set of AAA solutions that include the 802.1X client, various switches that support authentication, and the HyperBoss authentication and accounting system. The 802.1X client is installed on the user's PC. When the user needs to access the network, the user must authenticate with the 802.1X client, and only an authenticated user can use the network. The switch receives the authentication request from the client and transmits the user name and password to the HyperBoss authentication and accounting system. The switch itself does not do the actual authentication work. HyperBoss receives the authentication request sent by the switch, performs the actual authentication, and bills the user who is successfully authenticated.

The 802.1x protocol is used for communication between the 802.1x client and the switch, and the RADIUS protocol is used for communication between the switch and HyperBoss.

12.1 802.1x Introduction

The 802.1X protocol is a port-based access control and authentication protocol. The port here refers to a logical port, which may be a physical port, a MAC address, or a VLAN ID. The network switches implement the 802.1X protocol based on the MAC address.

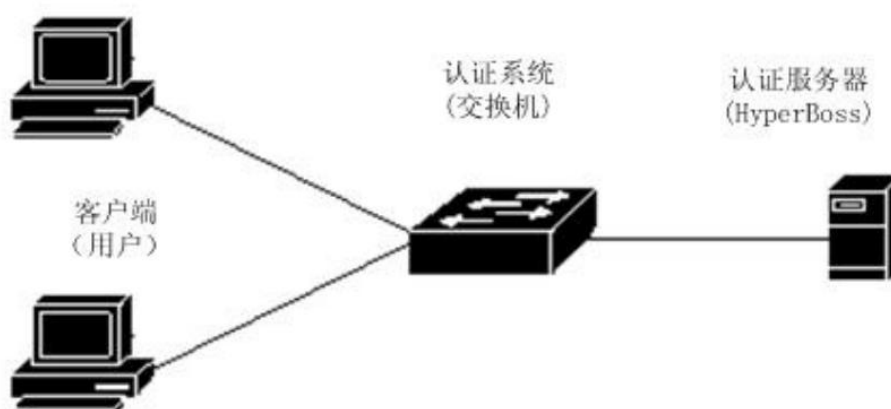
802.1X is a Layer 2 protocol. The authenticating switch and the user's PC must be in the same subnet, and the protocol packets cannot cross network segments. 802.1X authentication uses a client-server model, and there must be a server to authenticate all users. Before the user passes authentication, only the authentication flow can pass through the port of the switch. After authentication succeeds, the data flow can pass through the port of the switch. That is, the user must pass authentication before accessing the network.

This section includes the following topics:

- 802.1x device composition
- Introduction to Protocol Packets
- Protocol flow interaction
- 802.1x Port Status

12.1.1 802.1x Device Composition

An 802.1x device consists of three parts: a client (Supplicant System), an authentication system (Authenticator System), and an authentication server (Authentication Server System), as shown in the following figure.



The client is the device that requests access to the network, generally a user terminal system such as the user's PC. The user terminal system must have 802.1X client software installed, which implements the client part of the 802.1X protocol. The client initiates an 802.1X authentication request that asks the authentication server to authenticate its user name and password. If authentication succeeds, the user can access the network.

The authentication system is the device that performs authentication, such as a switch. The authentication system controls whether a user can access the network through the status of that user's logical port (which refers to the MAC address). If the status of the user's logical port is unauthorized, the user cannot access the network. If the status of the user's logical port is authorized, the user can access the network.

The authentication system is a relay between the client and the authentication server. The authentication system requests the identity information of the user, forwards that identity information to the authentication server, and forwards the authentication result sent by the authentication server to the client. The authentication system must implement the server part of the 802.1X protocol at the end close to the user, and the client part of the RADIUS protocol at the end close to the authentication server. The RADIUS protocol client of the authentication system packages the EAP information sent by the 802.1X client in RADIUS and sends it to the authentication server. It also decapsulates the EAP information from the RADIUS protocol packet sent by the authentication server and transmits it to the 802.1X client through the 802.1X server part.

The authentication server is the device that actually authenticates the user. The authentication server receives the identity information of the user sent by the authentication system and verifies it. If authentication succeeds, the authentication server authorizes the authentication system and allows the user to access the

network. If authentication fails, the authentication server tells the authentication system that user authentication failed and the user cannot access the network. The authentication server communicates with the authentication system through the RADIUS protocol extended by EAP. The network provides the HyperBoss authentication and accounting system to authenticate and account for users.

12.1.2 Introduction to Protocol Packets

The authentication data stream that the 802.1X protocol transmits on the network is in the format of an EAPOL (EAP Over LAN) frame. All user identity information, including the user name and password, is encapsulated in EAP (Extensible Authentication Protocol), and EAP is further encapsulated in an EAP frame. The user name is carried in EAP as clear text, and the password is carried in EAP as MD5 encryption.

The EAPOL frame format is shown in the following figure. PAE Ethernet Type is the Ethernet protocol type for EAPOL, with a value of 0x888E. Protocol Version is the EAPOL version number, with a value of 1. Packet Type is the EAPOL frame type. Packet Body Length is the length of the EAPOL frame content. Packet Body is the content of the EAPOL frame. The switch uses three types of EAPOL protocol frames:

	Octet Number
PAE Ethernet Type	1-2
Protocol Version	3
Packet Type	4
Packet Body Length	5-6
Packet Body	7-N

EAPOL-Start: the value of Packet Type is 1. This is the authentication initiation frame. When the user needs to be authenticated, the client sends this frame to the switch first.

EAPOL-Logoff: the value of Packet Type is 2. This exit request frame is sent to notify the switch when the user no longer needs to use the network.

EAP-Packet: the value of Packet Type is 0. This authentication information frame carries authentication information.

The format of the EAP packet is shown in the following figure. Code is the type of EAP packet: Request, Response, Success, or Failure. Identifier is an identifier used to match a Response with a Request. Length is the length of the EAP packet, including the header. Data is the EAP packet data.

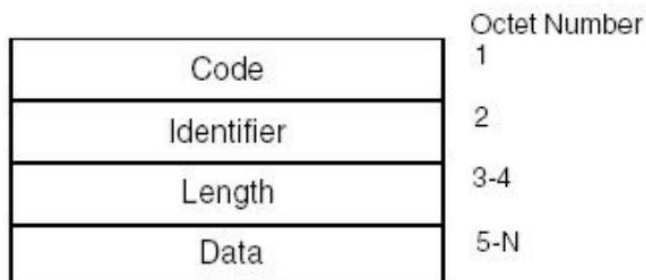
EAP packets include the following four types:

- a. EAP-Request: the value of Code is 1. The EAP request packet is sent from the switch to the client to request the user name and/or password.
- b. EAP-Response: the value of Code is 2. The EAP response packet is sent from the client to the switch, and it sends the user name and/or password to the switch.

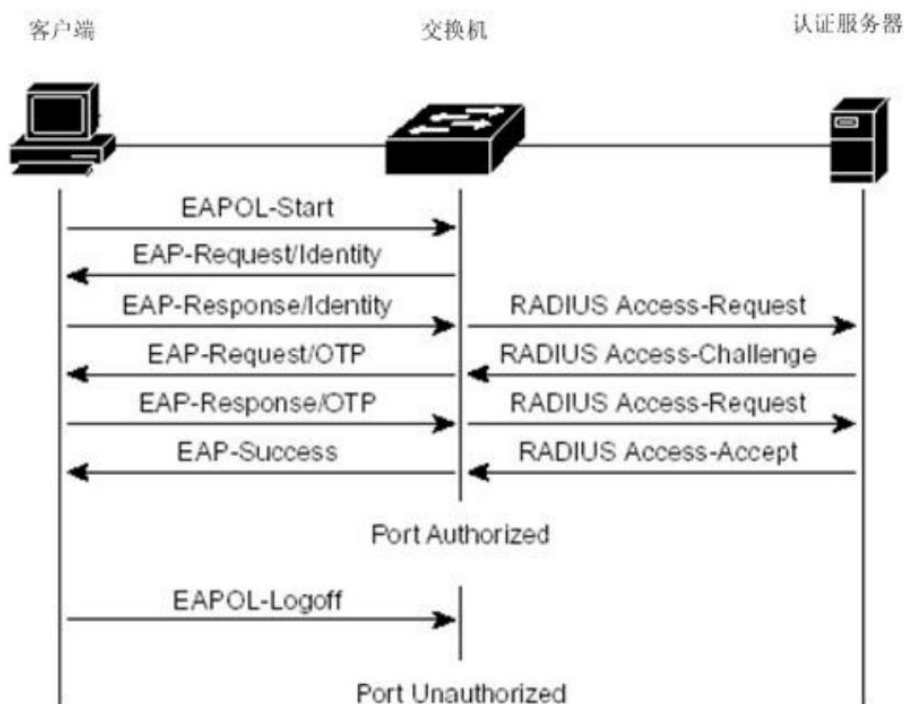
c. EAP-Success: the value of Code is 3. The EAP success packet is sent from the switch to the client to inform the client that user authentication succeeded.

d. EAP-Failure: the value of Code is 4. The EAP failure packet is sent from the switch to the client to inform the client that user authentication failed.

12.1.3 Protocol Flow Interaction



When 802.1x is enabled on the switch and the status of the port is Auto, all users accessing the port must be authenticated before they can access the network. The protocol interaction is shown in the figure below.



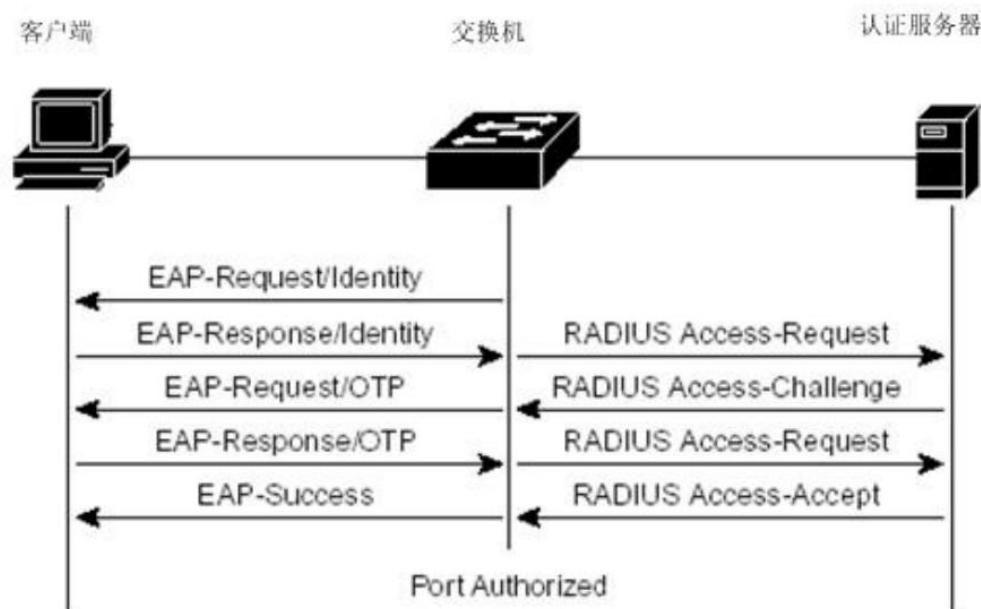
When a user needs to access the network, the client first sends an EAPOL-Start to the switch to request authentication. After receiving the authentication request, the switch sends an EAP-Request to ask for the user name. The client returns an EAP-Response, and the switch extracts the EAP information, encapsulates it in a RADIUS packet, and sends it to the authentication server. The authentication server requests the user's password, and the switch sends an EAP-Request to the client asking for the password. The client sends back an

EAP-Response, and the switch encapsulates the EAP information in a RADIUS packet and sends it to the authentication server. The authentication server authenticates the user according to the user name and password. If authentication succeeds, the authentication server notifies the switch, and the switch sends EAP-Success to the client and places the user's logical port in the authorized state. When the client receives EAP-Success, authentication has succeeded and the user can access the network.

When the user no longer needs to use the network, the client sends an EAPOL-Logoff to the switch, and the switch changes the state of the user's logical port to unauthorized. The user can no longer access the network.

To prevent a client from going offline abnormally, the switch provides a re-authentication mechanism. You can set the re-authentication interval on the switch. When the authentication time is up, the switch initiates re-authentication. If authentication succeeds, the user can continue to use the network. If authentication fails, the user cannot use the network. The protocol interaction is shown in the figure below.

12.1.4 802.1x Port Status



The port state referred to here is the physical port state of the switch. A physical port on a switch has four states: N/A, Auto, Force-authorized, and Force-unauthorized. When 802.1X is not enabled on the switch, all ports are in the N/A state. To set a switch port to the Auto state, the Force-authorized state, or the Force-unauthorized state, you must enable 802.1X on the switch.

When a port of the switch is in the N/A state, all users under the port can access the network without authentication. When the switch receives 802.1X protocol packets on the port, it discards them.

When a port of the switch is in the Force-authorized state, all users under the port can access the network without authentication. When the switch receives an EAPOL-Start packet on the port, it sends back an EAP-Success packet. When the switch receives other 802.1X protocol packets on the port, it discards them.

When a port of the switch is in the Force-unauthorized state, no user under the port can access the network at any time, and an authentication request can never pass. When the switch receives 802.1X protocol packets on the port, it discards them.

When a port of the switch is in the Auto state, all users under the port must be authenticated before accessing the network. The 802.1X protocol interaction is shown in the figure. If users need to be authenticated, set the port to the Auto state.

When a switch port is set to the Auto state, the anti-ARP spoofing function is also enabled. The anti-ARP spoofing function allows the port to forward only IP packets whose source MAC and source IP match the information the client provided during authentication, and only ARP packets whose sender IP and sender MAC match the information the client provided during authentication. Other packets are discarded. When you configure this function, the client must have a statically configured IP address. If the IP address is obtained dynamically through DHCP, you can enable the DHCP SNOOPING protocol to achieve this function. For a more detailed introduction, refer to the IP MAC binding configuration.

12.2 RADIUS Introduction

When a user is authenticated, the switch and the authentication server interact using the RADIUS protocol with EAP extension support. RADIUS uses a client/server model. The switch implements the RADIUS client, and the authentication server implements the RADIUS server.

To secure the interaction between the switch and the authentication server, and to prevent interaction with unauthorized switches or unauthorized authentication servers, the switch and the authentication server authenticate each other. Both must use the same key. When the switch or the authentication server sends a RADIUS protocol packet, it generates a message digest for the packet with an HMAC algorithm based on the key. When the switch or the authentication server receives a RADIUS protocol packet, it verifies the message digest using the key. If the verification passes, the packet is considered a valid RADIUS protocol packet. Otherwise, it is an invalid RADIUS protocol packet and is discarded.

This section includes the following topics:

- Introduction to Protocol Packets
- Protocol flow interaction
- User Authentication Method

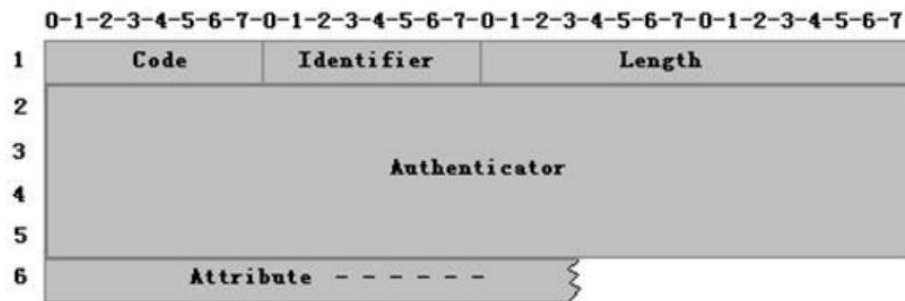
12.2.1 Introduction to Protocol Packets

RADIUS is a UDP-based protocol. It can encapsulate authentication information and accounting information. The early RADIUS authentication port was 1645; the current port is 1812. The early RADIUS accounting port was 1646; the current port is 1813.

Because RADIUS runs over UDP, it requires a timeout retransmission mechanism. To improve the reliability of communication between the authentication system and the RADIUS server, two RADIUS servers are generally used, that is, a standby server mechanism.

The format of a RADIUS message is shown in the figure below. Code is the RADIUS protocol message type. Identifier is used to match requests and replies. Length is the length of the whole message, including the

message header. Authenticator is a 16-byte string: a random number for a request packet, and an MD5-generated message digest for a reply packet. Attribute is an attribute in the RADIUS packet.



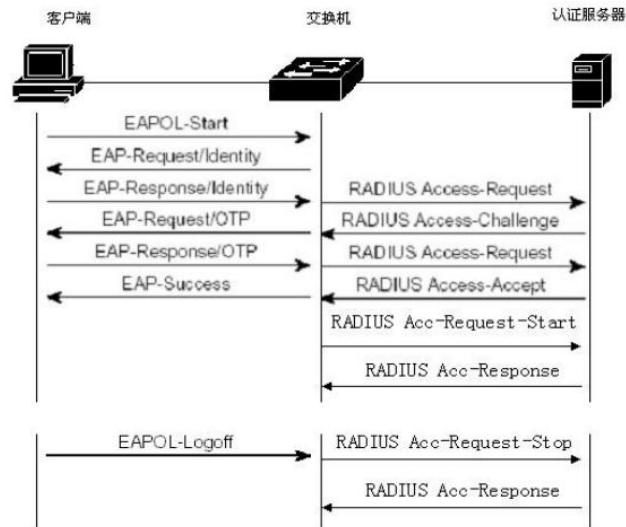
RADIUS message format

The network uses the following RADIUS protocol packets:

- Access-Request: The value of Code is 1. This is an authentication request packet sent from the authentication system to the authentication server. The user name and password are encapsulated in this packet.
- Access-Accept: The value of Code is 2. This is the response packet sent from the authentication server to the authentication system, indicating that user authentication succeeded.
- Access-Reject: The value of Code is 3. This is a response packet sent from the authentication server to the authentication system, indicating that user authentication failed.
- Access-Challenge: The value of Code is 11. This is a response packet sent from the authentication server to the authentication system, indicating that the authentication server requires further information from the user, such as a password.
- Accounting-Request: The value of Code is 4. This is the accounting request packet sent from the authentication system to the authentication server. It includes the start accounting packet and the end accounting packet. The accounting information is encapsulated in this packet.
- Accounting-Response: The value of Code is 5. This is an accounting response packet sent from the authentication server to the authentication system, indicating that the accounting information has been received.

12.2.2 Protocol Flow Interaction

When a user initiates authentication, the authentication system interacts with the authentication server through the RADIUS protocol. The following figure shows the protocol flow interaction in which the authentication system does not send RADIUS accounting packets. In general, after a user is authenticated successfully, or when a user goes offline, the authentication system must send a RADIUS accounting packet to the authentication server. The following figure shows that protocol flow interaction.



When a user is authenticated, the switch encapsulates the user name in an Access-Request message and sends the Access-Request message to the authentication server. The server responds with an Access-Challenge to request the user password. The switch requests the password from the client user, and the client encapsulates the password in the EAP. After acquiring the EAP, the switch encapsulates it in an Access-Request and sends the Access-Request to the authentication server. The authentication server authenticates the user. If the authentication succeeds, an Access-Accept is returned to the switch, and the switch notifies the client that the authentication succeeded after receiving the message. At the same time, an Accounting-Request is sent to notify the authentication server to start accounting, and the authentication server returns an Accounting-Response.

When the user no longer wants to use the network, the switch is notified that the user is offline. The switch sends an Accounting-Request to notify the authentication server to end the accounting. The accounting information is encapsulated in the packet, and the authentication server returns an Accounting-Response.

12.2.3 User Authentication Method

RADIUS has three methods of user authentication, as follows:

- PAP (Password Authentication Protocol)

The user passes the user name and password to the switch in clear text. The switch passes the user name and password to the RADIUS server in a RADIUS protocol packet. The RADIUS server searches the database. If the same user name and password exist, the authentication passes. Otherwise, the authentication does not pass.

- CHAP (Challenge Handshake Authentication Protocol)

When a user requests access to the Internet, the switch generates a 16-byte random code for the user. The user encrypts the random code, password, and other fields to generate a response, and transmits the user name and response to the switch. The switch passes the user name, the response, and the original 16-byte random code to the RADIUS server. The RADIUS server searches the database on the switch side according to the user name to obtain the same password used for encryption on the user side. It then encrypts the password

with the transmitted 16-byte random code and compares the result with the transmitted response. If they are the same, the authentication passes. If they are not the same, the authentication fails.

- EAP (Extensible Authentication Protocol)

With this authentication method, the switch does not really participate in the authentication. It only forwards traffic between the user and the RADIUS server. When the user requests access to the network, the switch requests the user name and transfers it to the RADIUS server. The RADIUS server generates a 16-byte random code for the user and stores the random code. The user encrypts the random code, password, and other fields to generate a response, and transmits the user name and response to the switch, and the switch forwards the response to the RADIUS server. The RADIUS server searches the database on the switch side according to the user name to obtain the same password as that used for encryption on the user side. It then encrypts the password with the stored 16-byte random code and compares the result with the transmitted response. If they are the same, the authentication passes. If they are not the same, the authentication fails. The authentication and accounting solution of the network adopts the EAP user authentication method.

12.3 Configuring 802.1x

This section describes the 802.1x configuration in detail. It includes the following topics:

- 802.1x Default Configuration
- Turning 802.1x On and Off
- Configure the 802.1x port state
- Configure the recertification mechanism
- Configure the maximum number of hosts connected to the port
- Configure the interval time and number of retransmissions
- Configure the port as a transmit port
- Configure the 802.1x client version number
- Configure whether to check the client version number
- Configure the authentication method
- Configure whether to check the timing packet of the client
- Displaying 802.1x information

12.3.1 802.1x Default Configuration

The default 802.1x configuration of the switch is as follows:

- 802.1x is off.
- The state of all ports is N/A.
- The recertification mechanism is off, and the recertification interval is 3600 seconds.
- The maximum number of access hosts for all ports is 100.
- The timeout interval for resending an EAP-Request is 30 seconds.
- The number of times an EAP-Request is retransmitted after a timeout is 3.

- The waiting time after a user authentication failure is 60 seconds.
- The timeout retransmission interval of the server is 10 seconds.

The switch provides a command in global CONFIG mode that returns all configurations to the default state, as follows:

```
Switch(config)#dot1x default
```

12.3.2 Turning 802.1x On and Off

The first step in configuring 802.1x is to enable it. In global CONFIG mode, enter the following command to enable 802.1x:

```
Switch(config)#dot1x
```

When 802.1x is turned off, all ports return to the N/A state. To turn off 802.1x, enter the following command in global CONFIG mode:

```
Switch(config)#no dot1x
```

12.3.3 Configure the 802.1x Port State

Be sure to enable 802.1x before you set the 802.1x port state. If all users on a port must be authenticated before they access the network, set the port to the Auto state.

The following command sets port ge1/1 to the Auto state in interface configuration mode and enables anti-ARP spoofing:

```
Switch(config-ge1/1)dot1x control auto
```

If the anti-ARP spoofing configuration fails, the cause may be one of the following:

- 1) System CFP resources are exhausted.
- 2) The current interface is configured with ACL filtering.
- 3) The current interface has the DHCP SNOOPING feature enabled. 4) The interface being configured is either a Layer 3 interface or a trunk interface.

The following command sets port ge1/1 to the Force-authorized state in interface configuration mode:

```
Switch(config-ge1/1)dot1x control force-authorized
```

The following command sets port ge1/1 to the Force-unauthorized state in interface configuration mode:

```
Switch(config-ge1/1)dot1x control force-unauthorized
```

The following command sets port ge1/1 to the N/A state in interface configuration mode:

```
Switch(config-ge1/1)no dot1x control
```

Note: If a port is bound to a MAC address, the port cannot be set to Auto, Force-authorized, or Force-unauthorized.

12.3.4 Configure the Recertification Mechanism

To prevent the switch and the authentication server from being unable to detect an abnormal offline of the client, the switch provides a re-authentication mechanism. The switch initiates an authentication every re-authentication interval.

The following command starts the recertification mechanism in global CONFIG mode:

```
Switch(config)#dot1x reauthenticate
```

The following command turns off the recertification mechanism in global CONFIG mode:

```
Switch(config)#no dot1x reauthenticate
```

The following command sets the recertification interval in global CONFIG mode:

```
Switch(config)#dot1x timeout re-authperiod <interval>
```

Note: Do not set the recertification interval too short. Otherwise, it consumes too much network bandwidth and too many CPU resources on the switch.

12.3.5 Configure the Maximum Number of Hosts Connected to a Port

Each port of the switch can control the maximum number of hosts that access it. This function prevents users from using multiple hosts to access the network without authorization. The maximum number of hosts connected to the port is 100 by default, and the maximum number can be set to 100. If the maximum number of hosts accessing a port is set to 0, the port denies any user access.

The following command sets the maximum number of hosts on port ge1/1 in interface configuration mode:

```
Switch(config-ge1/1)dot1x support-host <number>
```


12.3.6 Configure the Retransmission Interval and Count

The 802.1x protocol standard specifies some interval times and retransmission times for the protocol interaction and the protocol state machine. The switch uses the standard interval times and retransmission times. You are advised not to change these interval times or retransmission times when using the switch.

Tx-period represents the interval at which the switch resends the EAP-Request protocol packet. max-req represents the number of times the switch resends the EAP-Request. quiet-period represents the interval to wait for re-authentication when user authentication fails. Server-timeout indicates the interval at which the switch retransmits the RADIUS packet to the authentication server. supp-timeout indicates the interval at which the switch retransmits the EAP request packet to the client.

The following command configures these interval times and retransmission times in global CONFIG mode:

```
Switch(config)#dot1x timeout tx-period <interval>
Switch(config)#dot1x max-req <number>
Switch(config)#dot1x timeout quiet-period <interval>
Switch(config)#dot1x timeout server-timeout <interval>
Switch(config)#dot1x timeout supp-timeout <interval>
```

12.3.7 Configure the Port as a Transmit Port

When the switch does not enable 802.1x authentication but other switches in the subnet do, you can configure the port of the switch that connects the client and the authentication switch as the transmission port. The eapol authentication packet is then forwarded between the client and the 802.1x authentication switch, so that the other switches can perform 802.1x authentication of the client.

The following command sets port ge1/1 as a transmit port in interface configuration mode:

```
Switch(config-ge1/1)dot1x transmit-port
```

The following command sets port ge1/1 as a non-transmitting port in interface configuration mode:

```
Switch(config-ge1/1)no dot1x transmit-port
```

12.3.8 Configure the 802.1x Client Version Number

Configure the version number of the 802.1x client. Only a client whose version is not lower than the configured version number can be authenticated. Otherwise, the authentication fails. The default client version number for the switch is 2.0.

The following command configures the client version number in global CONFIG mode:

```
Switch(config)# dot1x client-version <string>
```

12.3.9 Configure Whether to Check the Client Version Number

Configure whether to check the version number of the 802.1x client. If checking is configured, the switch checks the client version number first when it authenticates the client. Checking is configured by default.

The following command turns on checking of the client version number in global CONFIG mode:

```
Switch(config)# dot1x check-version open
```

12.3.10 Configure the Authentication Method

Configure the authentication mode of the switch for 802.1x packets. The authentication mode initiated by the client is divided into general authentication and extended authentication. The switch can be configured to authenticate in the first place. If the authentication mode initiated by the client is inconsistent with the authentication mode configured on the switch, the client switches to the other authentication mode to initiate authentication after a certain number of authentication failures.

The following command configures the authentication mode of the switch as extended authentication in global CONFIG mode:

```
Switch(config)# dot1x extended
```

12.3.11 Configure Whether to Check the Client's Timing Packet

Configure whether the switch checks the timing packets of the client. After successful authentication, the switch requires the client to send 802.1x packets periodically, but not all clients send 802.1x packets periodically after they pass authentication. Use this command to configure whether the switch checks the timing packets of the client.

The following command is configured in global CONFIG mode for the switch to check the client's timing packets:

```
Switch(config)# dot1x check-client
```

12.3.12 Displaying 802.1x Information

The following command displays 802.1x information in normal mode or privileged mode. With show dot1x, all 802.1x configuration information is displayed, including the configuration information of all ports. With show dot1x interface, the information of all access users on the port is displayed:

```
Switch#show dot1x  
Switch#show dot1x interface
```

12.4 Configure RADIUS

This section describes the RADIUS configuration in detail. It includes the following topics:

- RADIUS Default Configuration
- Configure the IP address of the authentication server
- Configure Shared Secret
- Start and stop billing
- Configure RADIUS port and attribute information
- Configure RADIUS roaming
- Display RADIUS information

12.4.1 RADIUS Default Configuration

The default RADIUS configuration of the switch is as follows:

- The IP addresses of the primary and backup authentication servers are not configured, that is, the IP address is 0.0.0.0.
- No shared secret is configured, that is, the shared secret string is empty.
- Billing is enabled by default.
- The RADIUS authentication UDP port is 1812 and the billing UDP port is 1813.
- The RADIUS attribute NASPort has a value of 0xc353, NASPortType has a value of 0x0f, and NASPortServer has a value of 0x02.

12.4.2 Configure the Address of the Authentication Server

For RADIUS communication between the switch and the authentication server, the IP address of the authentication server must be configured on the switch. In practical applications, one authentication server or two authentication servers can be used: one as the primary authentication server and one as the backup authentication server. If the switch is configured with the IP addresses of two authentication servers, the switch can switch over to communicate with the backup authentication server when the communication between the switch and the primary authentication server is interrupted.

The following command configures the IP address of the primary authentication server in global CONFIG mode:

```
Switch(config)#radius-server host <ip-address>
```

The following command configures the IP address of the backup authentication server in global CONFIG mode:

```
Switch(config)#radius-server option-host <ip-address>
```

12.4.3 Configure Shared Secret

Mutual authentication is required between the switch and the authentication server, and the same shared key must be set on both the switch and the authentication server. Note that the shared key on the switch must be the same as the one on the authentication server.

The following command configures the shared secret of the switch in global CONFIG mode:

```
Switch(config)#radius-server key <string>
```

12.4.4 Start and Stop Billing

If accounting is turned off on the switch, the switch does not send RADIUS accounting packets to the authentication server after successful authentication or when the user goes offline. In general, billing is turned on in actual applications.

The following command starts accounting in global CONFIG mode:

```
Switch(config)#radius-server accounting
```

The following command turns off billing in global CONFIG mode:

```
Switch(config)#no radius-server accounting
```

12.4.5 Configure RADIUS Port and Attribute Information

You are advised not to modify the RADIUS port and attribute information configuration.

The following command modifies the RADIUS authentication UDP port in global CONFIG mode:

```
Switch(config)#radius-server udp-port <port-number>
```

The following command modifies RADIUS attribute information in global CONFIG mode:

```
Switch(config)#radius-server attribute nas-portnum <number>  
Switch(config)#radius-server attribute nas-porttype <number>  
Switch(config)#radius-server attribute service-type <number>
```

12.4.6 Configure RADIUS Roaming

When MAC, IP, or VLAN binding is performed on the client and the client is moved to another place, the bound client cannot perform 802.1x authentication, because its MAC address, IP address, or VLAN has changed. Turning on RADIUS roaming ignores the MAC, IP, or VLAN binding of the client and continues with 802.1x authentication.

The following command configures RADIUS roaming in global CONFIG mode:

```
Switch(config)#radius-server roam
```

The following command turns off RADIUS roaming in global CONFIG mode:

```
Switch(config)#no radius-server roam
```

12.4.7 Display RADIUS Information

The following command displays RADIUS configuration information in normal mode or privileged mode:

```
Switch#show radius-server
```

12.5 Configuration Examples

Turn on the 802.1x protocol, configure port ge1/1 as Auto, configure the primary authentication server as 198.168.80.111, and configure the shared key of the switch as abcdef.

```
Switch#  
Switch# dot1x  
Switch#config t  
Switch(config)#radius-server host 198.168.80.111  
Switch(config)#radius-server key abcdef  
Switch(config)# interface ge1/1  
Switch(config-ge1/1)# dot1x control auto
```

12.6 TACACS+ Introduction

TACACS+ authentication and authorization provides stricter user rights management. It can verify not only the legitimacy of users, but also authorize commands. After TACACS+ authentication is enabled, a user must have the user name and password verified through the TACACS+ server when accessing the switch. Only when the user name and password are correct and consistent can the user pass the authentication. Users can access the switch after authentication.

TACACS+ also divides user permissions into two levels: ordinary users and privileged users. Ordinary users can only stay in the normal mode of the CLI command line interface, and privileged users can access all modes of the CLI command line interface. The command execution permission is also set based on the permission level. When a user enters a command (except enable, end and exit), the permission must be verified on the TACACS+ server. If the verification fails, the command is not executed.

The TACACS+ authentication and authorization function applies only to telnet and SSH terminals. It does not control console terminals. When you access the switch through a telnet or SSH terminal, the user name and password must be verified. The CLI can be accessed only after the user name and password are verified. During SSH access, only privileged users can pass through. TACACS+ authentication also applies to web login, but it verifies only password privileges and permissions, without command authorization.

By default, the switch does not turn on the TACACS+ function. At this time, Telnet, SSH, and web login all use the multi-user management function. After the TACACS+ function is turned on, the multi-user management function can still be configured, but it is not actually used.

The commands related to TACACS+ authentication and authorization are shown in the table below:

Command	Description	CLI Mode
<code>tacacsplus enable</code>	Turn on the TACACS+ function	Global Configuration Mode
<code>tacacsplus disable</code>	Turn off the TACACS+ function	Global Configuration Mode
<code>tacacsplus host <server-ip></code>	Configure the primary server address. IPv4 and IPv6 are supported. Cisco ACS is recommended	Global Configuration Mode
<code>tacacsplus option-host <server-ip></code>	Configure the standby server address. IPv4 and IPv6 are supported. Cisco ACS is recommended	Global Configuration Mode
<code>tacacsplus key WORD</code>	Configure the shared key used to encrypt transmitted data. It must match the configuration on the server	Global Configuration Mode
<code>tacacsplus auth-type (PAP CHAP)</code>	Select the authentication method. The supported methods are PAP and CHAP. PAP is the default mode, and the field encapsulates the secret. CHAP encapsulates the MD5 check code of the secret.	Global Configuration Mode
<code>show tacacsplus</code>	View TACACS+ configuration information	Global Configuration Mode
<code>no tacacsplus host</code>	Clear the primary server address	Global Configuration Mode
<code>no tacacsplus key</code>	Clear the shared key	Global Configuration Mode

Chapter 13 GMRP Configuration

This chapter includes the following contents:

- Introduction to GMRP
- Configure GMRP
- Show GMRP

13.1 Introduction to GMRP

GMRP (GARP Multicast Registration Protocol) is a multicast registration protocol based on GARP. It is used to maintain multicast registration information in switches. All switches that support GMRP can receive multicast registration information from other switches, dynamically update the local multicast registration information, and propagate the local multicast registration information to other switches. This information exchange mechanism ensures the consistency of the multicast information maintained by all GMRP-supporting devices in the same switching network.

When a host wants to join a multicast group, it sends a GMRP join message. The switch adds the port that receives the GMRP join message to the multicast group, and broadcasts the GMRP join message in the VLAN where the receiving port is located, so that the multicast source in the VLAN can learn of the multicast member. When a multicast source sends a multicast packet to a multicast group, the switch forwards the multicast packet only to the ports connected to a member of the multicast group, thereby realizing Layer 2 multicast in the VLAN.

13.2 Configure GMRP

The main GMRP configurations include:

Open GMRP

View GMRP

In the configuration task, global GMRP must be enabled before port GMRP can be enabled.

13.2.1 Turn On GMRP Settings

Command	Description	Configuration Mode
<code>set gmrp enable disable</code>	Enable or disable global GMRP for all VLANs	Global Configuration Mode
<code>set gmrp enable vlan <vlan-id></code>	Enable global GMRP for a specific VLAN	Global Configuration Mode
<code>set gmrp registration{fixed forbidden normal} <if-name></code>	Configure the interface multicast registration mode	Global Configuration Mode
<code>set gmrp timer {join leave nleaveall} <time-value></code>	Configure the times of the various timers	Global Configuration Mode
<code>set port gmrp enable <if-name></code>	Enable the port GMRP function	Global Configuration Mode
<code>set port gmrp disable <if-name></code>	Disable the port GMRP function	Global Configuration Mode

13.2.2 Viewing GMRP Information

After you complete the configuration above, execute the show command in privileged mode to display the running status of GMRP. Verify the effect of the configuration by viewing the displayed information.

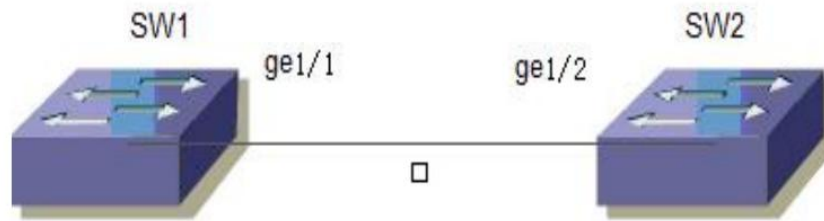
Command	Description	Configuration Mode
<code>show gmrp configuration</code>	View GMRP configuration information	Privileged Mode
<code>show gmrp machine</code>	View GMRP state machine information	Privileged Mode
<code>show gmrp statistics vlanid</code>	View GMRP statistics for a specific vlanid	Privileged Mode
<code>show gmrp timer <ifname></code>	View timer information for a specific port	Privileged Mode

13.3 GMRP Typical Configuration Example

(1) Networking requirements

To realize dynamic registration and update of multicast information between switches, GMRP must be started on the switch.

(2) Network diagram



GMRP Example Networking Diagram

Configuration steps:

Configure SW1

Start global GMRP

```
Switch(config)# set gmrp enable
```

Start port GMRP on Gigabit Ethernet port ge1/1

```
Switch(config)# set port gmrp enable ge1/1
Switch(config)#
```

Configure SW2

Start global GMRP

```
Switch(config)# set gmrp enable
```

Start port GMRP on Gigabit Ethernet port ge1/2

```
Switch(config)# set port gmrp enable ge1/2
Switch(config)#
```

Chapter 14 Configuring IGMP SNOOPING

In a MAN or the Internet, when the same data packet is sent by unicast to multiple receivers in the network, the packet must be copied to each receiving endpoint. The number of packets to be sent therefore increases linearly with the number of receivers, which increases the overall burden on the host, the switching and routing equipment, and network bandwidth resources. Efficiency is greatly affected. With the increasing demand for multipoint video conferencing, video on demand, and group communication applications, multicast has increasingly become a popular transmission mode in multipoint communication, because it improves resource utilization.

The switch implements the IGMP SNOOPING function to serve multicast applications. IGMP SNOOPING monitors the IGMP packets on the network to dynamically learn IP multicast MAC addresses.

This chapter describes the concept and configuration of IGMP SNOOPING. It includes the following topics:

- IGMP SNOOPING Introduction
- IGMP SNOOPING Configuration
- IGMP SNOOPING Configuration Example

14.1 IGMP SNOOPING Introduction

In a traditional network, multicast packets are treated as broadcast in a subnet, which easily causes network traffic and network congestion. When IGMP SNOOPING is implemented on the switch, it can dynamically learn IP multicast MAC addresses and maintain the output port list of each IP multicast MAC address, so that the multicast data stream is sent only to those output ports. This reduces network traffic.

This section includes the following topics:

- IGMP SNOOPING process
- Layer-2 dynamic multicast
- Join a group
- Leave a group

14.1.1 IGMP SNOOPING Process

IGMP SNOOPING is a Layer 2 network protocol. It monitors the IGMP packets that pass through the switch, maintains a multicast group according to the receiving port, VLAN ID, and multicast address of these IGMP packets, and then forwards these IGMP packets. Only the ports that have joined the multicast group receive the multicast data stream, which reduces network traffic and saves network bandwidth.

A multicast group includes the multicast group address, the member port, the VLAN ID, and the Age time.

The formation of an IGMP SNOOPING multicast group is a learning process. When a port of the switch receives an IGMP report packet, IGMP SNOOPING generates a new multicast group, and the port that received the IGMP report packet is added to the multicast group. When the switch receives an IGMP QUERY packet, if the multicast group already exists in the switch, the port that received the IGMP QUERY also joins the multicast

group. Otherwise, the IGMP QUERY packet is simply forwarded. IGMP SNOOPING also supports the Leave mechanism of IGMP V2. If IGMP SNOOPING is configured with fast-LEAVE as ENABLE, the receiving port can leave the multicast group immediately when it receives an IGMP V2 leave packet. If fast-leave-timeout is configured, the multicast group waits for this time to expire before leaving the multicast group.

IGMP SNOOPING has two update mechanisms. One is the leave mechanism described above. In most cases, IGMP SNOOPING uses the age time to delete expired multicast groups. When a multicast group joins IGMP SNOOPING, the joining time is recorded. When the multicast group remains in the switch for longer than the configured age time, the switch deletes the multicast group.

When a port receives a Leave protocol packet, the port is deleted from its multicast group immediately, which may affect the continuity of network data streams. A HUB or a network device without the IGMP SNOOPING function may be connected under the port, and many devices that receive multicast data streams may be connected under that device. If one device sends a Leave, it may affect the reception of the multicast data stream by other devices. The Fast-leave-timeout mechanism can prevent this. The Fast-leave-timeout mechanism configures a leave-waiting time. After receiving a leave packet, the port waits for the Fast-leave-timeout time before it is deleted from the multicast group to which it belongs. This makes it possible to guarantee the continuity of the network multicast stream.

14.1.2 Layer-2 Dynamic Multicast

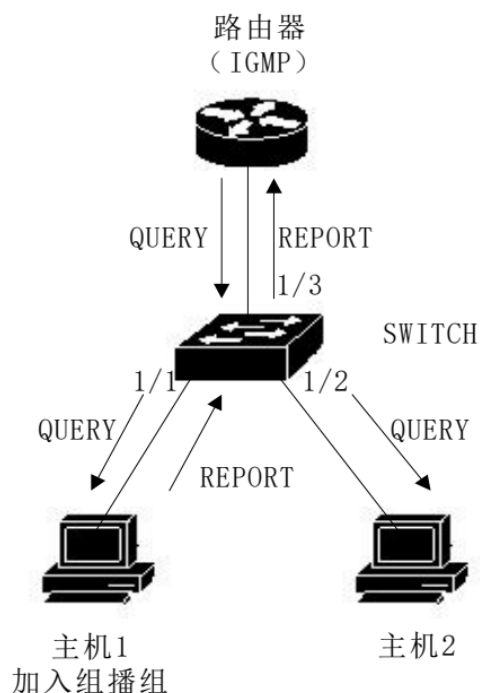
The multicast MAC address entries in the Layer 2 hardware multicast forwarding table can be learned dynamically by IGMP SNOOPING. The IP multicast MAC address is learned dynamically through IGMP SNOOPING.

When the switch turns off IGMP SNOOPING, the Layer 2 hardware multicast forwarding table is in unregistered forwarding mode. The multicast MAC address cannot be learned dynamically, there is no entry in the Layer 2 hardware multicast forwarding table, and all Layer 2 multicast data streams are treated as broadcasts.

When the network has a multicast environment, the switch can enable IGMP SNOOPING to control the multicast traffic of the network effectively. The Layer 2 hardware multicast forwarding table is then in register forwarding mode, and the switch can learn multicast MAC addresses by monitoring the IGMP protocol packets on the network. Only the Layer 2 multicast flows that match entries in the Layer 2 hardware multicast forwarding table can be forwarded.

14.1.3 Join a Group

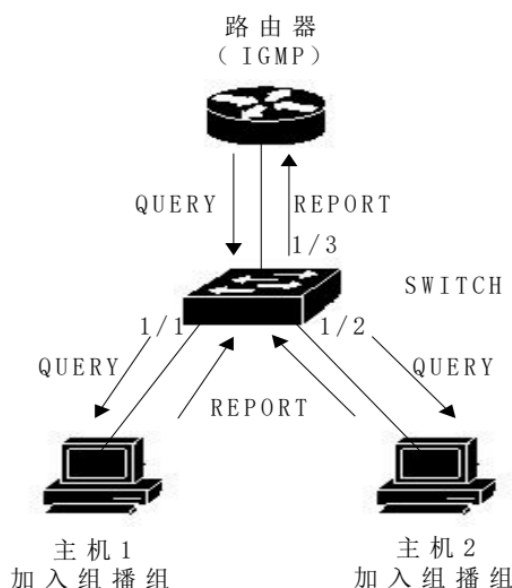
When a host wants to join a multicast group, the host sends an IGMP report packet that specifies the multicast group the host wants to join. When the switch receives an IGMP query packet, the switch forwards the packet to all other ports in the same VLAN. When the host under a port that wants to join the multicast group receives the IGMP query packet, it returns an IGMP report packet. When the switch receives the IGMP report packet, a Layer 2 multicast entry is created, and the port that received the IGMP QUERY packet and the port that received the IGMP report packet are added to the Layer 2 multicast entry as its output ports.



224.1.1.1

As shown in the figure above, all the devices are in one subnet, and the VLAN of the subnet is assumed to be 2. The router runs the IGMPv2 protocol and sends IGMP QUERY packets periodically. Host 1 wants to join the multicast group 224.1.1.1. When the switch receives an IGMP query packet on port 1/3, it notes this port and forwards the packet to ports 1/1 and 1/2. Host 1 sends back an IGMP report packet after receiving the IGMP query packet. Host 2 does not send an IGMP report packet, because it does not want to join the multicast group. When the switch receives the IGMP report packet on port 1/1, it forwards the packet out of query port 1/3 and creates a Layer 2 multicast entry (assuming the entry does not exist) that includes the following:

Layer 2 Multicast Address	VLAN ID	List of Output Ports
01:00:5e:01:01:01	2	1/1, 1/3



The condition in the above figure is the same as that in Figure 1. Host 1 has joined the multicast group 224.1.1.1, and now host 2 wants to join the multicast group 224.1.1.1. When host 2 receives the IGMP query packet and returns an IGMP report packet, the switch receives it on port 1/2, forwards it out of query port 1/3, and adds port 1/2 to the Layer 2 multicast entry. The Layer 2 multicast entry becomes:

Layer 2 Multicast Address	VLAN ID	List of Output Ports
01:00:5e:01:01:01	2	1/1, 1/2, 1/3

14.1.4 Leave a Group

To form a stable multicast environment, a device running IGMP, such as a router, sends an IGMP query packet to all hosts at regular intervals. A host that has joined or wants to join a multicast group receives the IGMP QUERY and sends back an IGMP REPORT.

If a host wants to leave a multicast group, there are two ways: active and passive. Active leave means that the host sends an IGMP leave packet to the router. Passive leave means that the host does not return an IGMP REPORT after receiving an IGMP QUERY from the router.

Corresponding to the ways in which a host leaves a multicast group, there are two ways for a port on the switch to leave a Layer 2 multicast entry: leaving by timeout, and leaving on receipt of an IGMP LEAVE packet.

When the switch does not receive an IGMP report packet of a multicast group on a port for a certain time, the port is removed from the corresponding Layer 2 multicast entry. If the Layer 2 multicast entry then has no port, the Layer 2 multicast entry is deleted.

When the fast-leave configuration of the switch is ENABLE, and a port receives an IGMP leave packet of a multicast group, the port is removed from the corresponding Layer 2 multicast entry. If the Layer 2 multicast entry then has no port, the Layer 2 multicast entry is deleted.

Fast-leave is generally used when a port is connected to a single host. If there is more than one host under a port, you can configure the fast-leave-timeout waiting time to ensure the continuity and reliability of multicast streams in the network.

14.1.5 IGMP Querier

In a network with a Layer 3 multicast device, the Layer 3 multicast device acts as the IGMP querier. The Layer 2 multicast device only needs to monitor the IGMP messages, and it can then establish and maintain the forwarding table entries to realize Layer 2 multicast. In a network without a Layer 3 multicast device, there is no Layer 3 multicast device to act as the IGMP querier. To enable the Layer 2 multicast device to monitor the IGMP messages, the IGMP querier function must be configured on the Layer 2 device. The Layer 2 multicast device then not only acts as the IGMP querier, but also monitors IGMP messages, so that it can establish and maintain forwarding table entries and realize Layer 2 multicast.

How it works:

With the IGMP querier function, the Layer 2 device plays the role of the IGMP routing querier. It sends IGMP query messages at regular intervals, intercepts and maintains the IGMP Report messages answered by users, and establishes the forwarding table entries of Layer 2 multicast. You can adjust the relevant parameters of the query messages sent by the IGMP querier through configuration.

1 Start the querier: You can enable the querier function on a specified VLAN.

2 Specify the IGMP version used by the querier: Specify the IGMP version used by the query messages sent by the querier. It can be configured as v1, v2, or v3.

3 Configure the source IP of the querier: configure the source IP address carried by the query messages sent by the querier

4 Configure the query interval of the querier: configure the time interval of the query messages sent by the global querier.

14.1.6 IGMP Snooping Multicast Filtering

A device running IGMP Snooping can control the multicast service range and load, and can effectively prevent unauthorized multicast streams. By configuring multicast filtering rules globally and applying the rules on the interfaces, you can allow or restrict the joining of specific groups.

14.2 IGMP SNOOPING Configuration

14.2.1 IGMP SNOOPING Default Configuration

IGMP SNOOPING is off by default, and the Layer 2 hardware multicast forwarding table is in unregistered forwarding mode.

Fast-leave is off by default.

Fast-leave-timeout is 300 seconds.

The default age time for the multicast group report port is 400 seconds.

The default age time for the multicast group QUERY port is 300 seconds.

14.2.2 Turn IGMP SNOOPING ON and OFF

The IGMP SNOOPING protocol can be turned on globally or separately for some VLANs. For a given VLAN, the IGMP SNOOPING protocol can only be turned on or off globally.

Turn on global IGMP SNOOPING

```
Switch#configure terminal
Switch(config)#ip igmp snooping
```

Turn on IGMP SNOOPING for one VLAN

```
Switch#configure terminal
Switch(config)#ip igmp snooping vlan <vlan-id>
```

Turn off global IGMP SNOOPING

```
Switch#configure terminal
Switch(config)#no ip igmp snooping
```

Turn off IGMP SNOOPING for one VLAN

```
Switch#configure terminal
Switch(config)#no ip igmp snooping vlan <vlan-id>
```

14.2.3 Configure the Lifetime

Configure the lifetime of a multicast group

```
Switch#configure terminal
Switch(config)#ip igmp snooping group-membership-timeout <interval> vlan <vlan-id>
```

Interval is in milliseconds.

Configure the lifetime of a query group

```
Switch#configure terminal
Switch(config)#ip igmp snooping query-membership-timeout <interval> vlan <vlan-id>
```

Interval is in milliseconds.

14.2.4 Configuring Fast-leave

Start fast-leave for one VLAN

```
Switch#configure terminal
Switch(config)#ip igmp snooping fast-leave vlan <vlan-id>
```

Close fast-leave

```
Switch#configure terminal
Switch(config)#no ip igmp snooping fast-leave vlan <vlan-id>
```

Configure the fast-leave wait times

```
Switch#configure terminal
Switch(config)# ip igmp snooping fast-leave-timeout <interval> vlan <vlan-id>
```

Restore the default fast-leave wait times

```
Switch#configure terminal
Switch(config)#no ip igmp snooping fast-leave-timeout vlan <vlan-id>
```

14.2.5 Configure MROUTER

Configure a static query port

```
Switch#configure terminal
Switch#interface ge1/6
Switch(config-ge1/6)#ip igmp snooping mrouter vlan [vlan-id]
```

14.2.6 Configure the IGMP Snooping Query Port Function

Configure a static query port

```
Switch#configure terminal
Switch(config)#interface ge1/6
Switch(config-ge1/6)#ip igmp snooping mrouter vlan [vlan-id]
```

14.2.7 Configuring the IGMP Snooping Query Feature

Start the IGMP snooping query function of vlan1


```
Switch#configure terminal
Switch(config)#ip igmp sno
Switch(config)#ip igmp snooping querier vlan 1
```

14.2.8 Configure IGMP Snooping Multicast Filtering

Configure the filtering multicast address of port ge1/1 as 235.0.0.1

```
Switch#configure terminal
Switch(config)#ip igmp snooping filter-rule 1 deny 235.0.0.1
Switch(config)#interface ge1/1
Switch(config-ge1/1)#ip igmp snooping filter-group 1
```

14.2.9 Display Information

Display IGMP SNOOPING configuration information

```
Switch#show ip igmp snooping
```

Display the configuration information of a VLAN

```
Switch#show ip igmp snooping vlan <vlan-id>
```

Display the aging information of a report multicast group

```
Switch#show ip igmp snooping age-table group-membership
```

Display the aging information of QUERY

```
Switch#show ip igmp snooping age-table query-membership
```

Display the forwarding information of a multicast group

```
Switch#show ip igmp snooping forwarding-table
```

Display MROUTER information

```
Switch#show ip igmp snooping mrouter
```

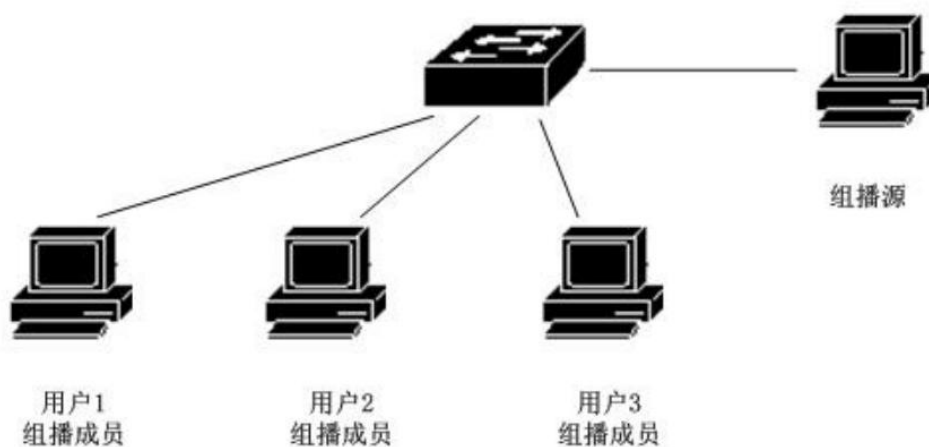
Display the current configuration of the system, including the configuration of IGMP SNOOPING

```
Switch#show running-config
```

14.3 IGMP SNOOPING Configuration Example

14.3.1 Configurations

Enable the IGMP SNOOPING function on the switch. User 1, User 2, and User 3 can join a specific multicast group.



```
Switch#config t
Switch(config)#ip igmp snooping
Switch(config)#vlan database
Switch(config-vlan)#vlan 200
Switch(config-vlan)#exit
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode access
Switch(config-ge1/1)#switchport access vlan 200
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 200
Switch(config)#interface ge1/3
Switch(config-ge1/3)#switchport mode access
Switch(config-ge1/3)#switchport access vlan 200
Switch(config)#ip igmp snooping vlan 200
Switch(config)#ip igmp snooping group-membership-timeout 60000 vlan 200
```

Chapter 15 MVR Configuration

This chapter includes the following contents:

- Introduction to MVR
- Configure MVR

15.1 Introduction to MVR

Multicast VLAN Registration (MVR) is used for multicast streaming applications in a service provider network, such as video on demand. MVR allows subscribers on a port to subscribe to or cancel multicast flows within a multicast VLAN, which allows data flows within one multicast VLAN to be shared by other VLANs. MVR has two objectives: (1) to deliver multicast streams between VLANs efficiently and securely with simple configuration; (2) to support dynamic joining and leaving of multicast groups.

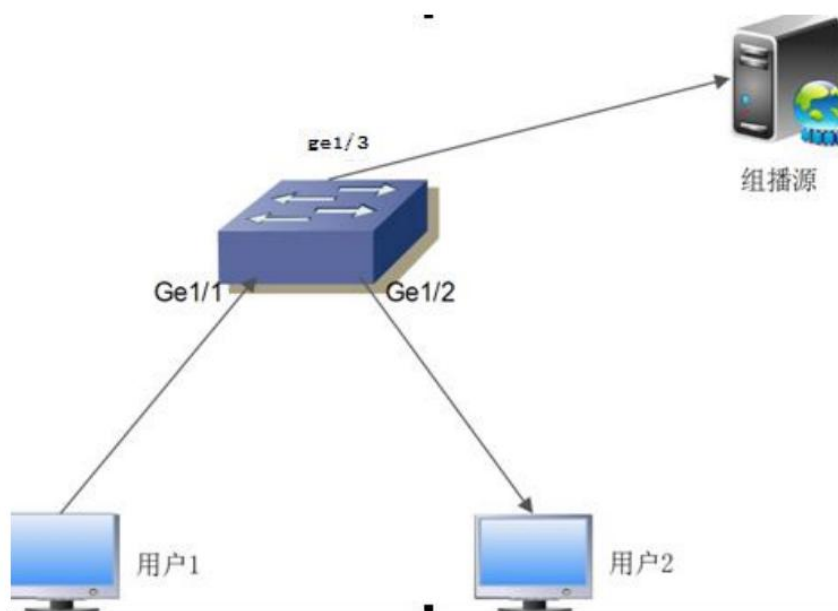
MVR operates in a way similar to IGMP snooping. The two functions can be started at the same time. MVR only handles the joining and leaving of the configured multicast group; the joining and leaving of other groups are managed by IGMP snooping. The difference between the two is that a multicast stream in IGMP snooping can only be forwarded in one VLAN, while a multicast stream in MVR can be forwarded in different VLANs.

15.2 Configure MVR

Command	Description	CLI Mode
<code>mvr (enable disable)</code>	Start global MVR	Global Configuration Mode
<code>no mvr</code>	Clear all MVR configuration	Global Configuration Mode
<code>mvr group A.B.C.D</code>	Configure the IP multicast address	Global Configuration Mode
<code>no mvr group A.B.C.D</code>	Delete the IP multicast address	Global Configuration Mode
<code>mvr group A.B.C.D <1-256></code>	Configure the IP multicast address and configure a continuous MVR group address	Global Configuration Mode
<code>mvr vlan <1-4094></code>	Specify the VLAN that receives the multicast data	Global Configuration Mode
<code>no mvr vlan</code>	Restore the default VLAN 1 for receiving multicast data	Global Configuration Mode
<code>mvr-interface (enable disable)</code>	Start interface MVR	Interface Configuration Mode
<code>show mvr</code>	Display MVR configuration information	Privileged Mode

15.3 MVR Configuration Example

The network topology is shown in the following figure. User 1 and user 2 belong to vlan10 and vlan20 respectively. User 1 and user 2 watch the same program, the program range is 225.1.1.1 ~ 225.1.1.64, and the MVR VLAN is 100:



Configure the VLAN, start global IGMP snooping, configure the MVR VLAN and the MVR program group scope, and enable MVR globally:

```
Switch#configure terminal
Switch(config)#ip igmp snooping
Switch(config)# mvr enable
Switch(config)#mvr vlan 100
Switch(config)#mvr group 225.1.1.1 64
Switch#
```

Configure the user ports ge1/1 and ge1/2, and the uplink port ge1/3 of the switch:

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode hybrid
Switch(config-ge1/1)# switchport hybrid native vlan 10
Switch(config-ge1/1)#switchport hybrid allowed vlan add 100 egress-tagged disable
Switch(config-ge1/1)#mvr enable
Switch(config-ge1/1)#
Switch#configure terminal
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode hybrid
Switch(config-ge1/2)# switchport hybrid native vlan 20
Switch(config-ge1/2)#switchport hybrid allowed vlan add 100 egress-tagged disable
Switch(config-ge1/2)#mvr enable
Switch(config-ge1/2)#
Switch#configure terminal
Switch(config)#interface ge1/3
Switch(config-ge1/3)# switchport access vlan 100
Switch(config-ge1/3)#
```

Chapter 16 Configuring DHCP SNOOPING

In a dynamic access network environment, a host acquires an IP address and network parameters through a DHCP server. DHCP SNOOPING is a snooping protocol for ARP attacks. By monitoring the DHCP messages, it dynamically binds the IP address assigned to the client by the DHCP server to the MAC address of the client, so that ARP attack messages can be filtered on the switch.

The switch supports the DHCP SNOOPING function, which can effectively defend against ARP attacks. DHCP SNOOPING intercepts DHCP messages on the network and binds port ARP information.

Four linked DHCP server physical ports can be configured to prevent unknown servers from interfering with the network to some extent.

This chapter describes the concept and configuration of DHCP SNOOPING. It includes the following topics:

- Introduction to DHCP SNOOPING
- DHCP SNOOPING Configuration
- DHCP SNOOPING Configuration Example

16.1 Introduction to DHCP SNOOPING

Because of its simple trust mechanism, the ARP protocol creates a loophole in network security. When an ARP attack packet carrying false MAC information arrives at a host, it directly overwrites the local ARP cache table without restriction, so that normal data flows to the attacker. Therefore, ARP information binding of the port is implemented on the Layer 2 switch, which can effectively filter ARP attack packets so that the attack packets cannot reach the attacked host. If an unpredictable DHCP server enters the network, IP address allocation becomes confused. The DHCP SNOOPING protocol provides binding of the physical port that links to the server. A non-designated physical port cannot forward the DHCP protocol packets sent by a DHCP server, which reduces the chance of such an unknown server entering the network.

This section includes the following topics:

- DHCP SNOOPING Handling Procedure
- DHCP SNOOPING Binding Table
- DHCP SNOOPING Bind Server's Physical Port

16.1.1 DHCP SNOOPING Handling Procedure

The DHCP SNOOPING protocol intercepts only DHCPrequest, DHCPack, and DHCPrelease messages. It does not receive other types of DHCP messages. It binds the mapping relationship between IP and MAC according to these messages.

The global DHCP SNOOPING switch is responsible for turning on the switch to receive DHCP messages, that is, IP messages with UDP ports 67 and 68.

16.1.2 DHCP SNOOPING Binding Table

The DHCP SNOOPING binding table entry is indexed by MAC address and contains the entry type, IP address, MAC address, interface information, delay timer, and lease timer. There are two types: REQ and ACK. An entry of the REQ type indicates that the DHCPrequest message has been received but the DHCPack message has not been received. At this time, the delay timer is started. The default time interval is 10 seconds. If no DHCPack message is received within 10 seconds, the binding table entry of the REQ type is deleted. An entry of the ACK type indicates that a DHCPack message has been received. The recorded IP address is the IP address allocated by the server, and a lease timer is started at this time. The time interval is the lease value provided by the DHCP server in the DHCPack message. The timer is restarted when the lease is renewed, and the binding table entry is deleted when the lease expires. The interface information records the interface where the client is located, that is, the interface corresponding to the binding relationship between the IP address and the MAC address.

When a DHCPrequest message is received, a binding table entry is created, the entry type is REQ, the IP address, the MAC address, and the interface information are recorded, and a 10-second delay timer is started.

When a DHCPrequest message is received and a binding table entry of the REQ type already exists, the entry is updated and the delay timer is restarted.

When a DHCPrequest message is received and a binding table entry of the ACK type already exists, the interface information is recorded.

When a DHCPack message is received, if a binding table entry of the REQ type exists, the IP address allocated by the server in the DHCPack message is recorded, the delay timer is closed, and the lease timer is started.

When a DHCPack message is received and no binding table entry of the REQ type exists, the message is discarded.

When a DHCPack message is received and a binding table entry of the ACK type already exists, if the interface has changed, the binding table entry of the original interface is deleted and the entry is updated.

If the interface has not changed but the IP address assigned by the server has changed, the binding table entry of the original interface is deleted and the entry is updated.

If the interface has not changed and the IP address has not changed, this indicates a renewal process, and the lease timer can be restarted.

The delay timer expires and the binding table entry of the REQ type is deleted.

The lease timer times out and the binding table entry of the ACK type is deleted.

16.1.3 Specify the Physical Port the DHCP Server Is Connected To

DHCP SNOOPING specifies the physical port of the linked server. DHCP messages can be received only on the specified port. If there are multiple DHCP servers in the network, the OFFER provided by a server on a non-designated port is filtered, and that server cannot assign an IP address to the client. The designated port supports unified allocation of IP addresses in the network. It avoids the case where the address pool of an unknown server is not in the IP plan and some clients cannot connect to the network normally. It also reduces the probability of abnormal network communication caused by private access to a server.

16.2 DHCP SNOOPING Configuration

16.2.1 DHCP SNOOPING Default Configuration

DHCP SNOOPING is off by default.

The default delay timer interval for entries of the REQ type in the DHCP SNOOPING binding table is 10 seconds.

16.2.2 Turning DHCP SNOOPING On and Off Globally

DHCP SNOOPING cannot be turned on or off for an interface until DHCP SNOOPING is turned on globally. You must turn it off on all interfaces before you turn off DHCP SNOOPING globally.

Turn on global DHCP SNOOPING

```
Switch#configure terminal
Switch(config)#ip dhcp snooping [IF_LIST]
```

The parameter is the physical port list of the linked DHCP server to be bound. Four ports can be specified in total. The port list is separated by ",", such as ge1/1, ge1/4, ge1/5

Turn off global DHCP SNOOPING

```
Switch#configure terminal
Switch(config)#no ip dhcp snooping
```

16.2.3 Interface DHCP SNOOPING ON and OFF

Turn on DHCP SNOOPING for a port

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#dhcp snooping
```

Turn off DHCP SNOOPING on one port

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#no dhcp snooping
```

16.2.4 Interface DHCP SNOOPING OPTION82 ON and OFF

Turn on DHCP SNOOPING OPTION82 for an interface


```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#dhcp snooping option82
```

Turn off DHCP SNOOPING OPTION82 on one port

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#no dhcp snooping option82
```

Configure the circuit-id of DHCP SNOOPING OPTION82 for a port

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)# dhcp snooping option82 circuit-id vlan111
```

Delete the circuit-id of DHCP SNOOPING OPTION82 for a port

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)# no dhcp snooping option82 circuit-id
```

16.2.5 Display Information

Display DHCP SNOOPING configuration information

```
Switch#show dhcp snooping
```

Display DHCP SNOOPING binding table information

```
Switch#show dhcp snooping binding-table
```

Display the current configuration of the system, including the DHCP SNOOPING configuration.

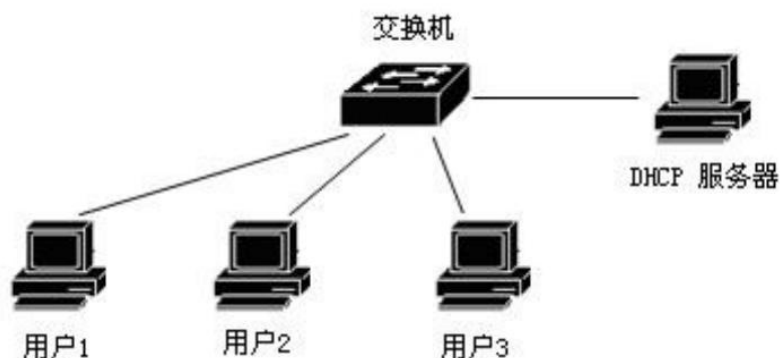
```
Switch#show running-config
```

16.3 DHCP SNOOPING Configuration Example

16.3.1 Configurations

Enable the DHCP SNOOPING function on the Layer 2 switch. User 1, User 2, and User 3 dynamically obtain their IP addresses and network parameters through the DHCP server. The DHCP SNOOPING OPTION82 function is

enabled on the interfaces of user 1, user 2, and user 3, the circuit-id is AAA, and ARP information is dynamically bound on the interfaces.



```
Switch#configure terminal
Switch(config)#ip dhcp snooping ge1/5
Switch(config)#interface ge1/1
Switch(config-ge1/1)#dhcp snooping
Switch(config-ge1/1)#dhcp snooping option82
Switch(config-ge1/1)#dhcp snooping option82 circuit-id aaa
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#dhcp snooping
Switch(config-ge1/2)#dhcp snooping option82
Switch(config-ge1/2)#dhcp snooping option82 circuit-id aaa
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#dhcp snooping
Switch(config-ge1/3)#dhcp snooping option82
Switch(config-ge1/3)#dhcp snooping option82 circuit-id aaa
Switch(config-ge1/3)#end
```

View DHCP snooping information

```
Switch#show dhcp snooping
```

DHCP Snooping is enabled globally

DHCP Server interface: ge1/5

Enable interface: ge1/1 ge1/2 ge1/3

Option 82 interface: ge1/1(Circuit ID: aaa) ge1/2(Circuit ID: aaa)

ge1/3(Circuit ID: aaa)

```
Switch#
Switch#show dhcp snooping binding-table
```

IP MAC FLAG PORT LEASE

192.168.1.100 00:11:5b:34:42:ad ACK ge1/1 23:59:58

192.168.1.101 00:11:64:52:13:5d ACK ge1/2 23:50:01

192.168.1.102 00:11:80:4d:a2:46 ACK ge1/3 20:34:45

16.4 DHCP SNOOPING Configuration Troubleshooting

If the DHCP snooping configuration fails, the cause may be one of the following:

- 1 The system CFP resource is exhausted.
- 2 If an interface is configured with the ACL filtering function, turning on DHCP SNOOPING globally fails.
- 3 If an interface is configured with IP and MAC binding, turning on DHCP SNOOPING globally fails.
- 4 The current interface is configured with the ACL filtering function.
- 5 802.1x anti-ARP spoofing is enabled on the current interface.
- 6 The configured interface is a Layer 3 interface or a trunk interface.

Chapter 17 DHCP CLIENT Configuration

17.1 Introduction to DHCP CLIENT

DHCP (Dynamic Host Configuration Protocol) is based on the Client/Server working mode. With the DHCP CLIENT function, the Layer 3 interface address of the switch can obtain its address and gateway through the DHCP server.

This section includes the following topics:

- Configuration of DHCP CLIENT

17.2 DHCP CLIENT Configuration

Turn on the DHCP client feature for interface vlan1

```
switch(config)#  
Switch(config)#interface vlan1  
Switch(config-vlan1)#dhcp client enable
```

Retrieve the IP address for interface vlan1

```
switch(config)#  
Switch(config)#interface vlan1  
Switch(config-vlan1)#dhcp client renew
```

Turn on the DHCP client feature for interface vlan1

```
switch(config)#  
Switch(config)#interface vlan1  
Switch(config-vlan1)#dhcp client release
```

Chapter 18 Configuring DHCP RELAY

This chapter includes the following contents:

- Introduction to DHCP RELAY
- DHCP RELAY Configuration
- DHCP RELAY Configuration Example

18.1 Introduction to DHCP RELAY

DHCP (Dynamic Host Configuration Protocol) is an enhanced version of BOOTP. It dynamically configures the network environment for the hosts on the network. It is divided into server and client. The server centrally manages IP network data, processes client requests, and dynamically configures the TCP/IP environment of the client. When DHCP works, there is at least one server on the network that can listen for DHCP requests from hosts on the network and negotiate TCP/IP parameters. Addresses can be allocated automatically or dynamically. In the automatic mode, once a client obtains an IP address, it uses that address permanently. In the dynamic mode, the IP address obtained by the client has a lease, and once the lease expires, the IP address must be released. The lease can also be renewed in advance, or another IP address can be leased. Dynamic allocation effectively solves the problem of insufficient IP addresses.

How DHCP works:

If the client is logging on to the network for the first time, it does not have any IP data and broadcasts a Discover message with source address 0.0.0.0 and destination address 255.255.255.255. If the server does not respond, four Discover requests are issued at regular intervals.

After receiving the Discover, the server selects an idle IP address and responds to the client with an Offer message.

If there are multiple servers on the network, the client receives multiple Offer messages. Generally, it selects the Offer that arrives first and broadcasts a Request message to tell all the servers which server's IP address it has accepted.

If the client finds through ARP that the IP address is already in use, it sends a Decline message to the server, rejects the Offer, and restarts the Discover process. After receiving the Request message, the server sends an Ack message to the client to confirm that the lease is effective.

If the client has already applied for a DHCP lease, the Discover process is generally not needed. Before the lease expires, the client uses the leased IP address to send a Request to the server for renewal. The server tries to let the client use the original IP address. If there is no problem, the server responds with an Ack message for confirmation. If the IP address is already used by another client, the server responds with a Nack message to reject the renewal request.

The client can use the Release message to actively release the lease.

The workstation sends a Request when it is powered on. It sends another Request when the lease is half expired, and the IP address can still be used if there is no confirmation. It sends a Request when the lease is 3/4 expired, and the IP address can no longer be used if there is no confirmation.

The Discover message is issued as a broadcast, so it can only stay in the same network segment; the router does not spread the broadcast message. When the server and the client are not in the same network segment, the client has not obtained its IP environment settings and does not know the location of the router, so the Discover message cannot reach the server. To solve this problem, you can use the DHCP relay function to let the router transfer the DHCP protocol messages, so that DHCP can operate across network segments.

18.2 DHCP RELAY Configuration

The DHCP relay function is mostly related to the interface. It forwards DHCP protocol messages across network segments, and the related configuration is performed in interface mode.

The configuration of DHCP-relay includes:

Start the DHCP-relay function of the interface

18.2.1 Start the DHCP-relay Function of the Interface

Mode: Interface configuration mode

Command: DHCP relay < ip-address-1 > ip-address-2] turns on the DHCP relay protocol on the interface

Command: no DHCP relay closes the DHCP relay protocol on the interface

Default: The DHCP relay protocol is not turned on.

18.2.2 Display Information

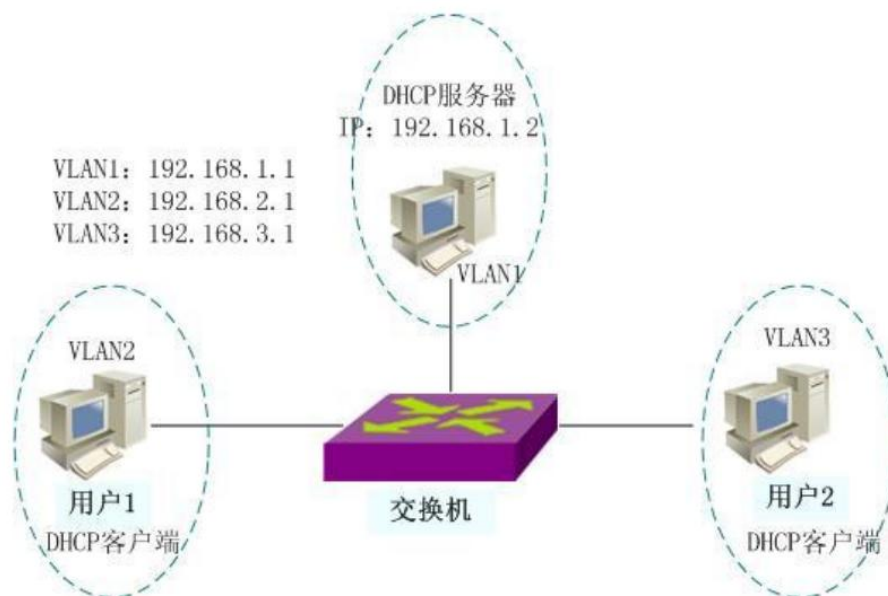
Display DHCP relay configuration information

```
Switch#show dhcp relay
```

18.3 DHCP RELAY Configuration Example

(1) Configuration

The switch must be configured for DHCP relay forwarding so that it can route and forward DHCP requests from User 1 and User 2 and DHCP reply acknowledgements from the DHCP server. User 1 and user 2 can then obtain valid IP addresses from DHCP servers in different network segments and access the network.



```
Switch>en
Switch# configure terminal
Switch(config)# vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
Switch(config-vlan)#exit
Switch(config)#ip interface vlan 2
Switch(config)#ip interface vlan 3
Switch(config)#int ge1/2
Switch(config-ge1/2)#sw access vlan 2
Switch(config-ge1/2)#int ge1/3
Switch(config-ge1/3)#sw access vlan 3
Switch(config-ge1/3)# interface vlan2
Switch(config-vlan2)#ip address 192.168.2.1/24
Switch(config-vlan2)#dhcp relay 192.168.1.2
Switch(config-vlan2)#interface vlan3
Switch(config-vlan3)#ip address 192.168.3.1/24
Switch(config-vlan3)#dhcp relay 192.168.1.2
```

(2) Verification

```
show running-config Display Configuration Command
```

show dhcp relay Displays DHCP relay configuration information

Chapter 19 Configuring DHCP SERVER

This chapter includes the following contents:

- Introduction to DHCP SERVER
- DHCP SERVER Configuration
- DHCP SERVER Configuration Example

19.1 Introduction to DHCP SERVER

DHCP (Dynamic Host Configuration Protocol) is an enhanced version of BOOTP. It dynamically configures the network environment for the hosts on the network. It is divided into server and client. The server centrally manages IP network data, processes client requests, and dynamically configures the TCP/IP environment of the client. When DHCP works, there is at least one server on the network that can listen for DHCP requests from hosts on the network and negotiate TCP/IP parameters. Addresses can be allocated automatically or dynamically. In the automatic mode, once a client obtains an IP address, it uses that address permanently. In the dynamic mode, the IP address obtained by the client has a lease, and once the lease expires, the IP address must be released. The lease can also be renewed in advance, or another IP address can be leased. Dynamic allocation effectively solves the problem of insufficient IP addresses.

How DHCP works:

If the client is logging on to the network for the first time, it does not have any IP data and broadcasts a Discover message with source address 0.0.0.0 and destination address 255.255.255.255. If the server does not respond, four Discover requests are issued at regular intervals.

After receiving the Discover, the server selects an idle IP address and responds to the client with an Offer message.

If there are multiple servers on the network, the client receives multiple Offer messages. Generally, it selects the Offer that arrives first and broadcasts a Request message to tell all the servers which server's IP address it has accepted.

If the client finds through ARP that the IP address is already in use, it sends a Decline message to the server, rejects the Offer, and restarts the Discover process. After receiving the Request message, the server sends an Ack message to the client to confirm that the lease is effective.

If the client has already applied for a DHCP lease, the Discover process is generally not needed. Before the lease expires, the client uses the leased IP address to send a Request to the server for renewal. The server tries to let the client use the original IP address. If there is no problem, the server responds with an Ack message for confirmation. If the IP address is already used by another client, the server responds with a Nack message to reject the renewal request.

The client can use the Release message to actively release the lease.

The workstation sends a Request when it is powered on. It sends another Request when the lease is half expired, and the IP address can still be used if there is no confirmation. It sends a Request when the lease is 3/4 expired, and the IP address can no longer be used if there is no confirmation.

The DHCP server protocol module receives Discover, Request, Decline, and Release messages. It dynamically allocates IP addresses to clients in the network and maintains its own address pool information and allocated client information.

19.2 DHCP SERVER Configuration

Configure DHCP server functions in global mode, interface mode, and address pool mode. Configuration includes startup commands, address pool configuration, global settings, and tuning switches.

DHCP server configuration includes:

- Start the global DHCP server function
- Start the interface to receive DHCP server messages
- Configure the address pool
- Configure the address pool range
- Configures the address pool subnet mask.
- Configures the address pool lease.
- Configures the address pool default gateway.
- Configures the address pool DNS server.
- Configures a manual exclusion address in the address pool.

19.2.1 Start the Global DHCP Server Function

Mode: Global Configuration Mode

Command: `ip dhcp server` starts the global DHCP server protocol

Command: `no ip dhcp server` turns off the global DHCP server protocol

Default: The DHCP server protocol is not enabled. Use this command to start the DHCP server protocol.

19.2.2 Enable the Interface to Receive DHCP Server Messages

Mode: Interface Configuration Mode

Command: `dhcp server listen`

The interface starts to receive DHCP server protocol messages.

Command: `no dhcp server listen`

The interface is closed and does not receive DHCP server protocol messages.

Default: The interface is not started and cannot receive DHCP server protocol messages.

19.2.3 Configure Address Pool

Mode: Global Configuration Mode

Command: `dhcp server pool <pool-name>`

Creates an address pool and enters address pool mode.

Command: `no dhcp server pool <pool-name>`

Deletes the specified address pool.

Parameter: `<pool-name>` is the address pool name, used to distinguish different address pools. Up to 16 characters.

Default: No address pool is configured. To configure an address pool, create an address pool name and enter address pool configuration mode; you do not need to configure the actual address.

19.2.4 Configure Address Pool Range

Mode: Address Pool Configuration Mode

Command: `no range` Deletes the address pool range

Parameters: `<low-address>` is the starting address of the pool range, in dotted-decimal format. `<high-address>` is the ending address of the pool range, in dotted-decimal format.

Default: No range is configured. When a range is configured in the address pool, an entry is created in the address pool for each dynamically assignable address within that range.

19.2.5 Configure Address Pool Subnet Mask

Mode: Address Pool Configuration Mode

Command: `subnet-mask <address>`

Configures the address pool subnet mask.

Parameters: `<address>` is the mask address, in dotted-decimal format. It is a variable-length mask.

Default: 255.255.255.0, a 24-bit mask.

19.2.6 Configure Address Pool Lease

Mode: Address Pool Configuration Mode

Command: `lease [<days> <hours> <minutes> | infinite]`

Configures the address pool lease.

Parameters: `<days>` is the number of days, in the range 0-999. `<hours>` is the number of hours, in the range 0-23. `<minutes>` is the number of minutes, in the range 0-59. All are integers. `infinite` means the lease is unlimited.

Default: The lease is 8 days.

19.2.7 Configure the Address Pool Default Gateway

Mode: Address Pool Configuration Mode

Command: default-router <ip-address>

Configures the address pool default gateway.

Command: no default-router Deletes the address pool default gateway

Parameters: <ip-address> is the default gateway IP address, in dotted-decimal format. It should be in the same network segment as the address pool range.

Default: The default gateway is not configured.

19.2.8 Configure Address Pool DNS Server

Mode: Address Pool Configuration Mode

Command: dns-server <ip-address1> [ip-address2]

Configures the address pool DNS server.

Command: no dns-server Removes the address pool DNS server

Parameters: <ip-address1> and <ip-address2> are the IP addresses of DNS servers, in dotted-decimal format. You can configure one or two DNS servers. If you configure two, enter them in one command, not in two separate commands. If you enter the command twice, the DNS server IP address entered later overrides the previously configured DNS server address, whether one or two DNS servers were previously configured.

Default: No DNS server is configured.

19.2.9 Configure Manually Excluded Addresses in the Address Pool

Mode: Address Pool Configuration Mode

Command: exclude-address <ip-address>

Configures a manual exclusion address in the address pool.

Command: exclude-address <low-address> <high-address>

Configures a manual exclusion address range.

Command: no exclude-address <ip-address>

Restores a manually excluded address.

Command: exclude-address <low-address> <high-address>

Configures manual exclusion address range recovery.

Command: no exclude-address all

Restores all manually excluded addresses in the address pool.

Parameters: <ip-address> is the manually excluded IP address, in dotted-decimal format. One command excludes one available address from the pool at a time. <low-address> and <high-address> are the starting and ending IP addresses of the manually excluded address range, in dotted-decimal format. One command can exclude multiple consecutive available addresses in the address pool range at a time. If an address in the range has already been manually excluded, it is skipped without any prompt. Manual exclusion means the address entry within the address pool range is not dynamically allocatable.

Default: No manual exclusion addresses are configured.

19.2.10 Configure OPTION82

Mode: Address Pool Configuration Mode

Command: option82 circuit-id <circuit-id>

Configures the circuit-id for option82.

Parameter: <circuit-id> is a string with a maximum length of 64.

19.2.11 Clear Allocated Address Table Entry

Mode: EXEC Configuration Mode

Command: clear dhcp server address {[ip-address] | [all]}

Deletes one allocated address table entry or all allocated address table entries.

19.2.12 Clear Address Entry with Conflict Detected

Mode: EXEC Configuration Mode

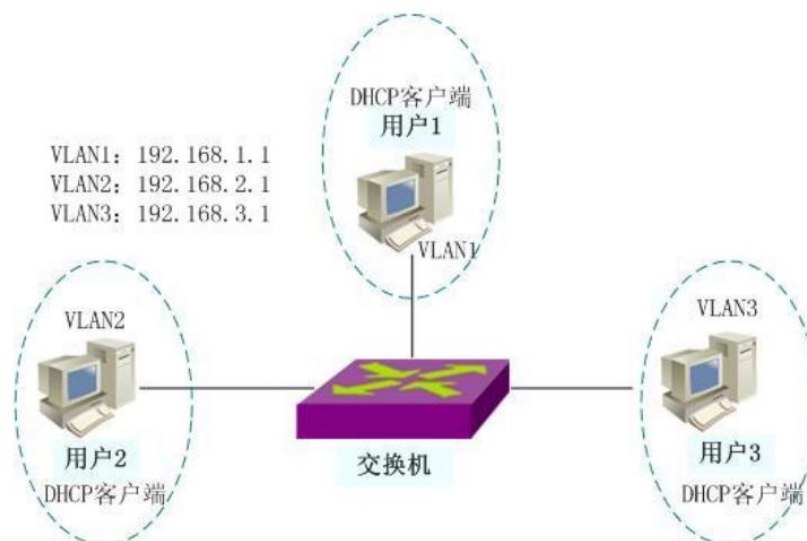
Command: clear dhcp server address conflict {[ip-address] | [all]}

Deletes addresses that are automatically excluded when a conflict is detected. You can delete all conflicting addresses, or a single address that was automatically excluded because of a conflict.

19.3 DHCP SERVER Configuration Example

(1) Configuration

Configure the corresponding address pools for the clients in the three different subnets vlan1, vlan2, and vlan3, so that the switch, acting as the DHCP server, can allocate IP addresses of the corresponding network segments to the clients in the three subnets.



```

Switch>en
Switch# configure terminal
Switch(config)#ip dhcp server
Switch(config)# vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
Switch(config-vlan)#exit
Switch(config)#ip interface vlan 2
Switch(config)#ip interface vlan 3
Switch(config)#int ge1/2
Switch(config-ge1/2)#sw access vlan 2
Switch(config-ge1/2)#int ge1/3
Switch(config-ge1/3)#sw access vlan 3
Switch(config-ge1/3)#interface vlan2
Switch(config-vlan2)#ip address 192.168.2.1/24
Switch(config-vlan2)#dhcp server listen
Switch(config-vlan2)#interface vlan3
Switch(config-vlan3)#ip address 192.168.3.1/24
Switch(config-vlan3)#dhcp server listen
Switch(config-vlan3)#interface vlan1
Switch(config-vlan1)#dhcp server listen
Switch(config)#dhcp server pool a
Switch(config-dhcp)#range 192.168.2.1 192.168.2.20
Switch(config-dhcp)#lease 2 0 0
Switch(config-dhcp)#default-router 192.168.2.1
Switch(config-dhcp)#dns-server 1.1.1.1 2.2.2.2
Switch(config-dhcp)#exclude-address 192.168.2.10
Switch(config-dhcp)#exit
Switch(config)# dhcp server pool b
Switch(config-dhcp)#range 192.168.3.2 192.168.3.20
Switch(config-dhcp)#default-router 192.168.3.1
Switch(config-dhcp)#exit
Switch(config)# dhcp server pool c
Switch(config-dhcp)#range 192.168.1.2 192.168.1.20
Switch(config-dhcp)#default-router 192.168.1.1
Switch(config-dhcp)#exit

```

(2) Verification

```
show running-config Display Configuration Command
```

show dhcp server Displays global configuration information

```
show dhcp server pool [pool-name]
```

Displays the address pool configuration information. Information for an individual address pool can be displayed.

```
show dhcp server address
```

Displays the assigned address table entry information.

Chapter 20 Configuring ACL

In a real network, the administrator is very concerned about network access security. The switch supports ACL filtering to provide network access security. Through the configuration of ACL rules, the switch filters the input data flow according to these rules to achieve network access security.

This chapter describes how to configure ACLs, including the following:

- ACL Resource Library Introduction
- ACL Filtering Introduction
- ACL Resource Pool Configuration
- ACL filtering configuration
- ACL Configuration Example

20.1 ACL Resource Library Introduction

The ACL (Access Control List) resource library is a collection of multiple groups of access rules. The ACL resource library does not control data forwarding itself; it is a collection of rules with conflict sorting. After an application references the ACL resource pool, that application controls the forwarding of data according to the rules provided by the ACL resource. ACLs can be applied to port access filtering, service access filtering, and QoS.

The ACL resource library has a standard IP rule group (group number 1 ~ 99, 1300 ~ 1999), an extended IP rule group (group number 100 ~ 199, 2000 ~ 2699), an IP MAC group <group number 700 ~ 799>, and an ARP group (number 1100 ~ 1199). Prioritization of conflicting rules is performed automatically within each set of rules. When you configure an ACL rule, the system inserts the rule at the appropriate location based on the collation.

During application, when a data packet passes through a port, the switch compares the field in each rule with all corresponding fields in the data packet. When multiple rules are completely matched at the same time, the first completely matched rule takes effect, and the matched rule determines whether the data packet is forwarded or discarded. An exact match is one in which the value of a field in the rule is exactly equal to the value of the corresponding field in the packet. A rule does not take the corresponding deny or permit action unless it exactly matches one of the ACL rules.

In a switch, rules within the same group are ordered automatically. The automatic sorting of rules is relatively complex. In the sorting process, rules with a large range are placed at the back, and rules with a small range are placed at the front. The size of the range is determined by the constraints of the rule. The fewer constraints a rule has, the larger the range of rule matches; the more constraints a rule has, the smaller the range of rule matches. The constraints of the rules are mainly reflected in the wildcard of the address and the number of some non-address fields. A wildcard is a bit string. The IP address is four bytes and the MAC address is six bytes. A bit of '1' means no match is required, and a bit of '0' means a match is required. Non-address fields refer to protocol type, IP protocol type, and protocol port. These fields also hide a wildcard. Their length is the byte length of the corresponding field, so the length of the same field is uniform and only the number of fields needs to be calculated. The more bits of the wildcard that are '0', the more constraints.

Take port access filtering as an example to illustrate the necessity of rule sorting and the advantages of automatic sorting. If you need to deny forwarding for the source address in the 10.10.10.0/16 network segment, and allow forwarding for the source address in the 192.168.1.0/24 network segment, you can configure the following two rules:

```
access-list 1 permit 192.168.1.0 0.0.0.255—Rule 1
access-list 1 deny 10.10.10.0 0.0.255.255 —Rule 2
```

Rule 1 and Rule 2 for short.

These two rules conflict, because the address of rule 1 is contained in the address of rule 2, and one is deny while the other is permit. According to the filtering principle of ACL, different orders give different results. To meet the above requirement, the order of the two rules must be: rule 1 first and rule 2 last. The switch implements this sorting automatically. No matter what order you configure the rules in, the final order is rule 1 in front of rule 2. When a packet with a source address of 192.168.1.1 is forwarded, the first rule is compared first, and then the second rule. If both rules match, the former rule takes effect (forwarding). If the source address is 10.10.10.1, only the first rule matches, so the packet is discarded (not forwarded).

Without sorting, you might configure rule 2 first and then rule 1, so rule 1 would be sorted later and rule 2 first.

```
access-list 1 deny 10.10.10.0 0.0.255.255—Rule 2
access-list 1 permit 192.168.1.0 0.0.0.255—Rule 1
```

Because rule 2, placed first, contains rule 1, placed later, a packet that exactly matches rule 1 may also exactly match rule 2, and rule 2 takes effect every time. This does not meet the requirements of the application.

In the switch, '0.0.255.255' is the wildcard bits, where a bit of '1' means no match is required and a bit of '0' means a match is required. The wildcard bits in rule 2 are '0.0.255.255', which needs to match two bytes (16 bits); the wildcard bits in rule 1 are '0.0.0.255', which needs to match three bytes (24 bits). So the rule scope of rule 2 is larger, and it is placed behind. In extended IP, more rule fields need to be considered for sorting, such as IP protocol type and communication port. Their ordering rules are the same: the more configuration restrictions, the smaller the scope of the rule, and vice versa. The sorting of rules is implemented in the background, and user commands can only be displayed in the order you configured them.

The filtering fields supported by ACL include source IP, destination IP, IP protocol type (such as TCP, UDP, OSPF), source port (such as 161), and destination port. You can configure different rules for access control according to different needs.

In a switch, one set of rules can be applied by multiple applications. For example, a set of rules can be referenced by both port access filtering and service access filtering, or by two port access filters.

20.2 ACL Filtering Introduction

ACL filtering is carried out at the input port of the switch. Port filtering is achieved by performing rule matching on the data flow input to the port. ACL filtering is handled at the line speed of the switch and does not affect the forwarding efficiency of the data flow.

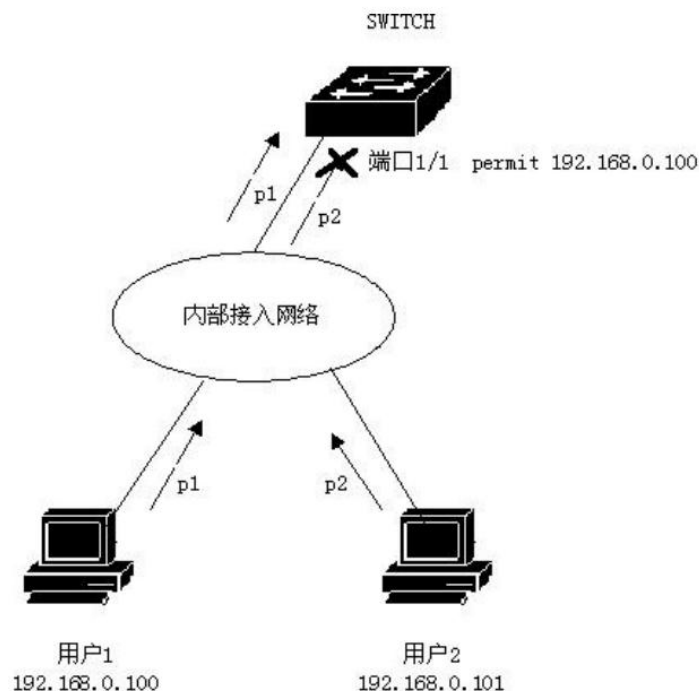
When a port of a switch is not configured with ACL filtering, all incoming traffic through the port does not match any rules and can be forwarded through the port. When a port of a switch is configured with ACL filtering, all input data flows passing through the port are matched against the rules. If the action of the matched rule is permit, the data flow is allowed to be forwarded. If it is deny, the data flow is not allowed to be forwarded and is discarded.

When you configure ACL filtering on a port, you can select multiple ACL rule groups for that port. After selection, the group of rules is imported into the CFP of the port. If the group of rules contains no rule that denies or allows all IP protocol packets, a rule that denies all IP protocols is added when the CFP is written. When the rules in the ACL repository change, the rules written to the CFP change automatically as well.

For example, if there is only one rule in a set of rules, `access-list 1 permit 192.168.1.0 0.0.0.255`, a rule that denies all IP protocol packets is hidden by default, so two rules are actually imported into the CFP of the port. In traffic filtering, only traffic with a source address from 192.168.1.0 to 192.168.1.255 can be forwarded through that port, and all other traffic is filtered out.

For example, there are two rules in the set: `access-list 1 deny 192.168.1.0 0.0.0.255` and `access-list 1 permit any`. Because one rule allows all IP protocol packets, there is no hidden rule, and two rules are actually imported into the CFP of the port. In traffic filtering, only traffic with a source address from 192.168.1.0 to 192.168.1.255 is filtered out, and all other traffic is allowed to be forwarded.

The following figure shows an example of ACL filtering. Port 1/1 of the switch selects ACL rule group 1, which has only one rule, `access-list 1 permit 10.10.10.100`. Two users under port 1/1 on the switch want to access the network from that port. The IP address of user 1 is 10.10.10.100 and the IP address of user 2 is 10.10.10.101. Only user 1 can access the network through port 1/1 of the switch. User 2 cannot. Data flow P1 from user 1 can be forwarded through port 1/1 of the switch, while data flow P2 from user 2 is dropped at port 1/1 of the switch.



When multiple ports are used for ACL filtering, the same ACL rule group can be selected and the same filtering rule can be used.

Whether a port references one set of rules or multiple sets, the rules are sorted automatically, even if the sorting between two sets of rules is crossed.

After you reference a set of rules, if that set of rules changes, the port referencing it responds to your configuration automatically. You do not need to reconfigure the reference on the port.

20.3 ACL Resource Pool Configuration

By default, the switch has no rules.

The resource pool in the switch supports four types of ACL rules: standard IP rules, extended IP rules, IP MAC groups, and ARP groups. ACL configuration is described by these four categories of rules.

Standard IP rules: Standard IP rules control packet forwarding by source IP address.

Command form: `access-list <groupId> {deny | permit} <source>`

Parameter description:

groupId: The access control list group number. Standard IP ACLs support groups from 1 to 99 or 1300 to 1999.

deny/permit: If there is an exact match, the packet is denied or allowed to be forwarded.

deny/permit: If there is an exact match, the packet is denied or allowed to be forwarded.

source: The source IP can be entered in three ways:

- 1) A.B.C.D wildcard can control IP addresses from one network segment;
- 2) any equals A.B.C.D 255.255.255.255
- 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0

Wildcard: Determines which bits need to be matched. '0' means a match is required and '1' means no match is required.

Extended IP rules: Extended IP rules are extensions to standard IP rules. They control packet forwarding by source IP, destination IP, IP protocol type, and service port.

Command form: `access-list <groupId> {deny | permit} <protocol>`

`<source> [eq <srcPort>] <destination> [destPort] <tcp-flag>`

Parameter description:

groupId: Access control list group number. Extended IP ACLs support groups from 100 to 199 or 2000 to 2699.

deny/permit: If there is an exact match, the packet is denied or allowed to be forwarded.

protocol: For protocol types above the IP layer, such as TCP or UDP, you can also enter the corresponding number, for example 6 (TCP). If you do not want to control these protocols, enter IP or 0.

source: The source IP can be entered in three ways:

- 1) A.B.C.D wildcard can control IP addresses from one network segment;
 - 2) any equals A.B.C.D 255.255.255.255
 - 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0
- srcPort: If the protocol is TCP or UDP, the source port of the data packet can be controlled. You can enter a familiar port service name, such as www, or a number, such as 80.

There are three ways to input the destination IP:

- 1) A.B.C.D wildcard can control IP addresses from one network segment;
 - 2) any equals A.B.C.D 255.255.255.255
 - 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0
- destPort: If the protocol is TCP or UDP, the destination port of the data packet can be controlled. The input method is the same as for srcPort.

tcp-flag: For the case where the protocol is TCP. You can control the TCP field matching of the packet. The optional parameters are ack, fin, psh, rst, syn, and urg.

IP MAC rules: IP MAC groups can control the source and destination MAC addresses and the source and destination IP addresses of IP packets.

Command form: `access-list <groupId> {deny | permit} <src-mac> ip`

`<src-ip> <dst-ip>`

Parameter description:

groupid: Access control list group number. Extended IP ACLs support groups from 700 to 799.

deny/permit: If there is an exact match, the packet is denied or allowed to be forwarded.

src-mac: Source MAC address.

There are three ways to enter a MAC address:

- 1) The HHHH.HHHH.HHHH wildcard can control the MAC address from one segment;
- 2) any is equivalent to HHHH.HHHH.HHHH FFFF.FFFF.FFFF.
- 3) host A.B.C.D is equivalent to HHHH.HHHH.HHHH 0000.0000.0000

dst-ip: Destination IP address.

There are three ways to enter an IP address:

- 1) A.B.C.D wildcard can control IP addresses from one network segment;
- 2) any equals A.B.C.D 255.255.255.255
- 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0

ARP rules: The ARP group can control the operation type, sender MAC, and sender IP of the ARP packet.

Command form: access-list <groupid> {deny | permit} arp {arp-type}

<sender-mac> <sender-ip>

Parameter description:

groupid: The access control list group number. Extended IP ACLs support groups from 1100 to 1199.

deny/permit: If there is an exact match, the packet is denied or allowed to be forwarded.

arp-type: Refers to any, reply, or request. any is the type that does not control the ARP packet, reply controls the ARP response packet, and request controls the ARP request packet.

The MAC address of the sender of the ARP packet.

There are three ways to enter a MAC address:

- 1) The HHHH.HHHH.HHHH wildcard can control the MAC address from one segment;
- 2) any is equivalent to HHHH.HHHH.HHHH FFFF.FFFF.FFFF.
- 3) host A.B.C.D is equivalent to HHHH.HHHH.HHHH 0000.0000.0000 IP address of the sender of the ARP packet.

There are three ways to enter an IP address:

- 1) A.B.C.D wildcard can control IP addresses from one network segment;
- 2) any equals A.B.C.D 255.255.255.255
- 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0

List of additional commands:

```
show access-list [groupId]
```

Displays a list of the rules configured in the current ACL. If groupId is entered, a list of rules for that group is displayed; otherwise, a list of all rules is displayed.

```
no access-list <groupId>
```

Deletes the specified rule list, that is, all rules for the groupId group.

20.4 ACL Time-Based ACL

A time period describes a particular time range. You may need some ACL rules to take effect at one or more specific times and not be used for message filtering at other times, commonly known as filtering by time period. In this case, configure one or more time periods first, then refer to a time period by its name under the corresponding rule. That rule takes effect only in the specified time period, which achieves ACL filtering based on time period.

If the time period referenced by the rule is not configured, the system gives a prompt and still allows the rule to be created successfully. However, the rule cannot take effect until you configure the referenced time period and the system time is within the specified time period.

There are two ways to configure the time period:

- (1) Configure a relative time period: from a certain hour and minute of a certain day to a certain hour and minute;
- (2) Configure an absolute time period: from a certain hour and minute on a certain day of a certain month in a certain year to a certain hour and minute on a certain day of a certain month in a certain year. To configure a time-based ACL:

Command	Description	CLI Mode
time-range WORD cycle-time from <0-23> <0-59> to <0-23> <0-59>	Configure a relative time period containing only hours and minutes for the time period	Global Configuration Mode
time-range WORD cycle-time days from <0-6> to <0-6>	Configure a relative time period containing only the day of the week for the time period X	Global Configuration Mode
time-range WORD cycle-time from <0-23> <0-59> to <0-23> <0-59> days from <0-6> to <0-6>	Configure a relative time period including time of day and week for time period X	Global Configuration Mode
time-range WORD utter-time from <2000-2100> <1-12> <1-31> <0-23> <0-59> to <2000-2100> <1-12> <1-31> <0-23> <0-59>	Configure an absolute time period containing year, month, day, and hour for time period X	Global Configuration Mode

Command	Description	CLI Mode
<code>no time-range WORD cycle-time</code>	Delete all relative time periods of a certain time period	Global Configuration Mode
<code>no time-range WORD utter-time</code>	Delete all absolute time periods for a certain time period	Global Configuration Mode
<code>no time-range WORD</code>	Delete a time period (including all relative and absolute time periods)	Global Configuration Mode
<code>no time-range</code>	Delete all time periods	Global Configuration Mode
<code>show time-range WORD cycle-time</code>	Display all relative time periods for a certain time period	Privileged Mode
<code>show time-range WORD utter-time</code>	Display all absolute time periods for a certain time period	Privileged Mode
<code>show time-range WORD</code>	Display a time period (including all absolute and relative time periods)	Privileged Mode
<code>show time-range</code>	Show all time periods	Privileged Mode
<code>time-acl (<1-99> <100-199> <1300-1999> <2000-2699> <700-799> <1100-1199>) time-range WORD</code>	Apply the given ACL rule for a certain time period; it takes effect when the ACL is applied on the interface	Global Configuration Mode
<code>no time-acl (<1-99> <100-199> <1300-1999> <2000-2699> <700-799> <1100-1199>) time-range (WORD)</code>	Unapply an ACL rule from a time period or from all time periods	Global Configuration Mode
<code>show time-acl (<1-99> <100-199> <1300-1999> <2000-2699> <700-799> <1100-1199>) time-range</code>	Displays all time periods to which the specified ACL rule applies	Privileged Mode
<code>show time-acl all time-range</code>	Displays the time period to which all ACL rules apply	Privileged Mode

Note the following:

- (1) If multiple relative time periods are configured for a certain time period, the relationship between the relative time periods is OR. The time period is active when the system time falls in any relative time period;
- (2) If multiple absolute time periods are configured for a certain time period, the absolute time periods are in an OR relationship. The time period is active when the system time falls in any absolute time period;
- (3) If a relative time period and an absolute time period are configured for a certain time period at the same time, the relative time period and the absolute time period are in an AND relationship. The time period is active only when the system time falls in both the relative and absolute time periods;

(4) A maximum of 256 time periods can be defined. A maximum of 256 relative and absolute time periods can be configured for one time period. An ACL rule can apply up to 256 time periods. When an ACL rule associated with a time period is applied to an interface, the time period takes effect.

20.5 ACL Filtering Configuration

By default, no port on the switch has ACL filtering.

Command list:

```
access-group <groupId>
```

Mode: Layer 2 Interface Configuration Mode

Parameters:

groupId: ACL group number bound to the port

Function: Configures ACL port filtering.

Note: If the above command fails or does not take effect, the reasons may be:

There are too many rules in the ACL group, or the hardware resources are exhausted or occupied by other applications.

Displays the ACL port filtering configuration

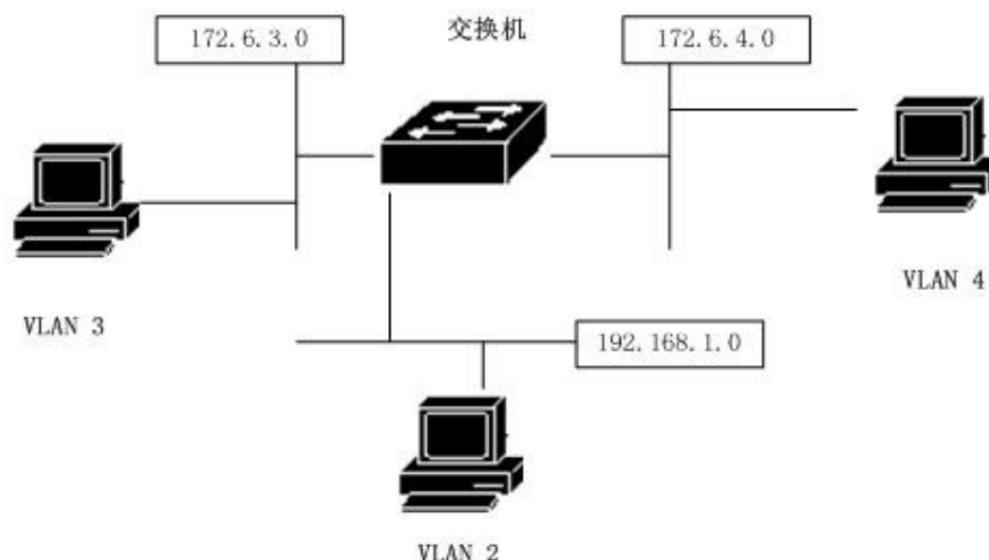
```
show access-group
```

Deletes the ACL port filtering configuration related to the current port

```
no acl- group <groupId>
```

20.6 ACL Configuration Example

A switch is connected to three subnets. An ACL is designed to block the source address of the 192.168.1.0 network address, while traffic from other network addresses is allowed. The 192.168.1.0 segment is connected to port 1/1 of the switch.



Configure the following on the switch:

```

Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#vlan 3
Switch(config)#interface ge1/1
Switch(config-ge1/1)# switchport mode access
Switch(config-ge1/1)#switchport access vlan 2
Switch(config)#interface vlan2
Switch(config-vlan2)#ip add 192.168.1.1/24
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 3
Switch(config)#interface vlan3
Switch(config-vlan2)#ip add 172.16.3.1/24
Switch(config)#access-list 10 deny 192.168.1.0 0.0.0.255
Switch(config)#access-list 10 permit any
Switch(config)#interface ge1/1
Switch(config-ge1/1)#access-group 10
Switch(config-ge1/1)#exit
Switch(config)#interface ge1/2
Switch(config-ge1/2)#access-group 10

```

20.7 ACL Configuration Troubleshooting

If the ACL configuration fails, the reasons may be:

(1) Make sure that all IPs are connected before configuring the access control list, then add the access control list. This access list blocks IP traffic with a source address in the 192.168.1.0 segment from passing through the switch. Note how the subnet complement is written. Use the show access-list command to list the access control list for viewing, and be sure that the source and destination addresses are not written backwards. Then view the access control list. The default access control list has an implicit deny any statement at the end. If you want other traffic to pass, add a permit any statement; otherwise it will not pass.

- (2) The system is configured with static IP MAC binding.
- (3) The DHCP SNOOPING protocol is enabled on the current interface.
- (4) The system CFP resource is exhausted.

Chapter 21 Basic TCP/IP Configuration

For a Layer 2 switch with a network management function, you must provide basic network configuration for the TCP/IP protocol to enable communication with other devices.

This chapter includes the following contents:

- Configure VLAN interface
- Configure ARP
- Configure Static Route
- IP Routing Configuration Example

21.1 Configure VLAN Interface

In a switch, each Layer 3 interface is attached to a VLAN, so Layer 3 interfaces are also called VLAN interfaces. VLAN interfaces are created and deleted manually. Up to 4094 VLANs can be created on the switch, but only up to 32 subnets can be established. You can create the subnet interface according to your requirements. You can delete the subnet interface manually, and it is also deleted when the VLAN where the subnet is located is deleted.

Each VLAN interface has a name. The name of a VLAN interface is the string "VLAN" followed by the VLAN ID number. For example, the name of the Layer 3 interface of VLAN 1 is "vlan1", and the name of the Layer 3 interface of VLAN 4094 is "vlan4094".

Like ports, VLAN interfaces have an administrative status and a link status. At present, the switch does not provide configuration of the management state of the VLAN interface. As long as the VLAN interface is created, its management state is always UP. The link state of a VLAN interface is related to the ports included in the VLAN corresponding to the interface. As long as the link state of one port in the VLAN is RUNNING, the link state of the VLAN interface is RUNNING. If no port in the VLAN is RUNNING, the link state of that VLAN interface is also not RUNNING.

On a VLAN interface, you can configure an IP address and indicate the network prefix (which can be translated to a netmask) for the segment connected to this interface. Currently, the switch supports only one IP address on one VLAN interface. Before configuring the IP address, create a VLAN and add the relevant port to the VLAN. By default, the switch has a VLAN1 interface with the IP address 10.10.10.1/24 set on it. You can also modify the IP address of the VLAN1 interface. Interfaces for VLANs other than VLAN 1 do not have IP addresses set by default.

The commands to configure the IP address of a VLAN interface are shown in the following table:

Command	Description	CLI Mode
<code>Ip interface vlan <2-4094></code>	Create a VLAN interface	Global Configuration Mode
<code>No Ip interface vlan <2-4094></code>	Delete a VLAN interface	Global Configuration Mode
<code>ip address <ip-prefix></code>	Set the IP address on the VLAN interface. The parameters include the IP address of the interface and the network prefix of the connected network segment. If the VLAN interface already has an IP address, the original IP address is deleted first, then the specified IP address is set. The format of the parameter is A.B.C.D/M.	Interface Configuration Mode
<code>no ip address [ip-prefix]</code>	Delete the IP address of the VLAN interface. If a parameter is specified, it must be the same as the parameter given during setup, or the command is invalid. The format of the parameter is A.B.C.D/M.	Interface Configuration Mode

The commands for viewing the VLAN interface are shown in the following table:

Command	Description	CLI Mode
<code>show interface [if-name]</code>	View the information of a VLAN interface, including IP address, MAC address, management status, and link status. The parameter is the interface name of the VLAN interface. If no parameter is specified, the information of all ports and VLAN interfaces is viewed.	Normal Mode, Privileged Mode
<code>show running-config</code>	View the current configuration of the system, including the configuration of the VLAN interface.	Privileged Mode

Example:

Configure the subnet 193.1.1.0 on the VLAN3 interface with a subnet prefix of 24 (that is, a mask of 255.255.255.0) and an interface IP address of 193.1.1.1, then view the information for the VLAN3 interface. The commands are as follows:

```
switch(config)#interface vlan3
switch(config-vlan3)#ip address 193.1.1.1/24
switch(config-vlan3)#end
switch#show interface vlan3
```

21.2 Configure ARP

ARP (Address Resolution Protocol) is a protocol that provides a mapping from an IP address to the corresponding MAC address. When a source sends an Ethernet data frame to a destination in the same VLAN, the destination is determined based on the 48-bit Ethernet MAC address, and the destination decides whether to receive the packet based on the destination MAC address of the packet.

Assume that hosts A and B in two adjacent network segments communicate through a switch. Before sending data to host B, host A first sends an ARP request message to the switch interface directly connected to host A, and then sends the data packet to that interface after receiving an ARP response. After receiving the data packet, the switch first broadcasts an ARP request message to host B. After receiving the ARP response message from host B, the switch sends the data packet to host B.

The switch has an ARP cache, called an ARP table, that holds a record of the mapping of IP addresses to MAC addresses on directly connected networks. Each entry in the ARP table has a lifetime, which is 20 minutes by default. If the switch does not receive an ARP request or response message from the IP address during the lifetime, the ARP entry corresponding to that IP address is deleted.

This section includes the following topics:

- Configure static ARP
- Configure ARP Binding
- View ARP information

21.2.1 Configure Static ARP

There are two kinds of ARP entries in the ARP table: static ARP and dynamic ARP. Static ARP is an ARP entry you configure through commands. The system does not refresh or delete it automatically; you must do so manually. Dynamic ARP is ARP that the system learns automatically from received ARP request or response packets. The system creates and deletes these entries automatically and updates and maintains them in real time without user intervention, but you can delete a dynamic ARP entry manually.

The switch has no static ARP entry configured by default. Note that when a VLAN interface is deleted, or the IP of the subnet segment of the interface is changed, the static and dynamic ARP entries in the original subnet segment are deleted.

The commands for configuring static ARP are shown in the following table:

Command	Description	CLI Mode
<code>arp <ip-address> <mac-address></code>	Configure a static ARP entry. The first parameter is the IP address, which must be within a subnet segment. The second parameter is the MAC address, which must be a unicast MAC address. The format of the MAC address is HHHH.HHHH.HHHH, such as 0010.5CB1.7825.	Global Configuration Mode
<code>no arp <ip-address></code>	Deletes an ARP entry, including the ARP entry for an IP address	Global Configuration Mode

21.2.2 View ARP Information

The commands for viewing ARP information are shown in the following table:

Command	Description	CLI Mode
<code>show arp</code>	View the ARP entry information in the ARP table, including all ARP entries.	Normal Mode, Privileged Mode
<code>show running-config</code>	View the current configuration of the system, including the ARP configuration.	Privileged Mode

21.3 Configure Static Route

A static route is a user-defined route that makes a packet traverse a specified path from a source address to a destination address. Packets that cannot be routed can be sent to the default gateway by configuring a static route as the default route.

Static routes are configured manually by an administrator. They suit networks with a simple structure, where the administrator only needs to configure the static route to make the switch work normally. Static routes do not consume valuable network bandwidth because there are no routing updates.

The default route is also a static route. Simply put, a default route is used only when no matching route entry is found. That is, the default route is used only when there is no suitable route. In the routing table, the default route appears as a route to the network 0.0.0.0/0 with a mask of 0.0.0.0. If the destination of the message is not in the routing table and there is no default route in the routing table, the message is discarded and an ICMP message indicating that the destination address or network is unreachable is returned to the source. A default route is very useful in a network. In a typical network with hundreds of switches, running a dynamic routing protocol may consume a large amount of bandwidth resources. Using a default route saves the time taken by routing and the bandwidth resources taken by packet forwarding, which to some extent can meet the needs of a large number of users communicating at the same time.

A switch can be configured with multiple static routes to the same destination, but only one of these routes is activated for actual data forwarding. The switch has no static routes configured by default.

The commands for configuring a static route are shown in the following table:

Command	Description	CLI Mode
<code>ip route <ip-prefix> <nexthop-address></code>	Set up a static route. The first parameter specifies the segment IP and network prefix length, and the second parameter specifies the next-hop IP address.	Global Configuration Mode
<code>ip route <ip-address> <mask-address> <nexthop-address></code>	The function is the same as the previous command. The first parameter specifies the IP address of the network segment, the second specifies the mask of the network segment, and the third specifies a next-hop IP address.	Global Configuration Mode
<code>no ip route <ip-prefix> [nexthop-address]</code>	Removes a static route. The first parameter specifies the network segment IP and network prefix length, and the second parameter specifies the next-hop IP address. Without the second parameter, all routes matching the specified network segment are removed. With the second parameter, the route that matches both the specified network segment and the next hop is removed.	Global Configuration Mode
<code>no ip route <ip-address> <mask-address> [nexthop-address]</code>	The function is the same as the previous command. The first parameter specifies the IP address of the network segment, the second specifies the mask of the network segment, and the third specifies the next-hop IP address. Without the third parameter, all routes matching the specified network segment are removed. With the third parameter, the route that matches both the specified network segment and the next hop is removed.	Global Configuration Mode

The commands for viewing the route are shown in the following table:

Command	Description	CLI Mode
<code>show ip route [<ip-address> <ip-prefix>]</code>	View the information of the activated route. You can view all routes, one route, the routes of a network segment, or the static routes.	Normal Mode, Privileged Mode
<code>show ip route database</code>	View information about all routes, both active and inactive, by selecting View All Routes.	Normal Mode, Privileged Mode
<code>show running-config</code>	View the current configuration of the system to see the static route configuration.	Privileged Mode

Example:

Set the destination network to 200.1.1.0, the subnet mask to 255.255.255.0, and the next hop to 10.1.1.2. The configuration commands are:

```
Switch(config)#ip route 200.1.1.0 255.255.255.0 10.1.1.2
```

or Switch(config)#ip route 200.1.1.0/24 10.1.1.2

Delete a static route with a destination IP address of 200.1.1.0, a subnet mask of 255.255.255.0, and a next hop of 10.1.1.2. The configuration commands are:

```
Switch(config)#no ip route 200.1.1.0/24
```

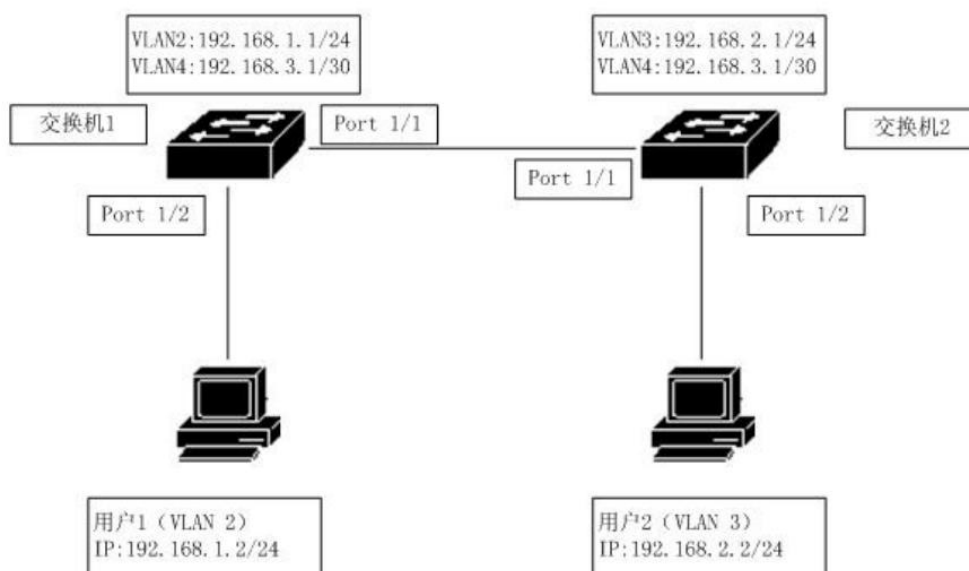
Or Switch(config)#no ip route 200.1.1.0/24 10.1.1.2

Or Switch(config)#no ip route 200.1.1.0 255.255.255.0

Or Switch(config)#no ip route 200.1.1.0 255.255.255.0 10.1.1.2

21.4 TCP/IP Basic Configuration Example

21.4.1 Three-Layer Interface



Configure the Layer 3 interface for VLAN2 on switch 1 and assign the IP address 192.168.1.1/24.

Configure it as follows:

```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#exit
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switch access vlan 2
Switch(config)#ip interface vlan 2
Switch(config)#interface vlan2
Switch(config-vlan2)#ip add 192.168.1.1/24
```

Verify that User1 can ping the Layer 3 interface IP address for VLAN 2 on switch 1.

21.4.2 Static Route

If user 2 wants to access switch 1, the traffic must go through the routing function of switch 2.

The configuration on switch 1 is as follows:

```
Switch#config t
Switch(config)#ip route 192.168.2.0 255.255.255.0 192.168.3.2
```

The configuration on switch 2 is as follows:

```
Switch#config t
Switch(config)#ip route 192.168.1.0/24 192.168.3.1
```

Verify that user 2 can ping switch 1.

21.4.3 ARP

Configure static ARP for user 1 to allow only user 1 to access from VLAN 2. Assume that the MAC address of user 1 is 00:00:00:01.

Switch 1 is configured as follows:

```
Switch#config t
Switch(config)#arp 192.168.1.2 0000.0000.0001
```

Verify that User1 can ping the Layer 3 interface IP address for VLAN 2 on switch 1.

Chapter 22 Configuring SNMP

The switch provides SNMP for remote management of the switch. This chapter describes how to configure SNMP, including the following:

This chapter includes the following contents:

- Introduction to SNMP
- SNMP Configuration
- SNMP Configuration Example

22.1 Introduction to SNMP

SNMP is a simple network management protocol and the most widely used network management protocol. It has five functions: fault management, accounting management, configuration management, performance management, and security management. It provides the information format for communication between network management application software and the network management agent.

The SNMP network management protocol has four elements: management station, management agent, management information base, and network management protocol. On the switch, the management agent is the server through which the management workstation accesses the switch. The information that the management workstation accesses on the network management agent is organized in the form of MIB to form a management information base.

SNMP has three operations: GET, SET, and TRAP. The GET operation lets the management station get the value of an object in the agent. The SET operation lets the management station set the value of an object in the agent. The TRAP operation lets the agent advertise events to the management station.

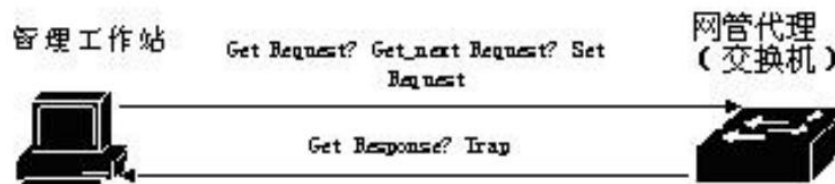
TRAP messages are sent to the management workstation when an event occurs on the switch. These messages include cold start, warm start, link up and link down of the port, authentication failure of the community name, and status switching of STP.

At present, there are three versions of SNMP: SNMPV1, SNMPV2, and SNMPV3. Each later version is an upgrade of the previous one, with enhanced functions and improved security. The switch supports all three versions of SNMP and can parse SNMP protocol packets of all three versions. A TRAP message may be sent using SNMPV1, SNMPV2, or SNMPV3.

The switch supports RFC, BRIDGE, and proprietary MIB objects, and can be fully managed via SNMP. Some of the MIBs supported by the switch are listed below: RFC 1213, RFC 1493, RFC 1724, RFC 1850, RFC 1907, RFC 2233, RFC 2571, RFC 2572, RFC 2573, RFC 2574, MIBs common to RFC 2575, RFC 2674, etc.

The figure shows an example of SNMP protocol interaction between a management workstation and a management agent. The management station can access the switch management agent by sending SNMP messages for Get Request, GetNext Request, GetBulk Request, and Set Request to obtain or set the values of the MIB objects of the switch. The switch management agent sends back an SNMP Get Response message to

the management workstation. When certain events occur on the switch, the management agent of the switch actively sends SNMP TRAP messages to the management station.



22.2 SNMP Configuration

SNMP configuration includes the community configuration of the switch, the TRAP workstation, and SNMP system information. By default, the switch has a read-only community named public. The switch can be configured with up to eight communities. The switch has no TRAP workstation configured by default.

SNMP commands are shown in the following table:

Command	Description	CLI Mode
<code>snmp community <community-name> {ro rw}</code>	Configure the community name used to access network management. This is an interactive command. During configuration, enter the name of the created community and the read/write permission as prompted.	Global Configuration Mode
<code>no snmp community <community-name></code>	Deletes the specified SNMP community name.	Global Configuration Mode
<code>snmp trap <notify-name> host <ipaddress> version {1 2c 3}</code>	Add or modify the destination to which the SNMP trap is sent. This is an interactive command. The notify name is unique. If you modify an existing name, the trap sending target item is modified. The host is the destination address for sending the trap; the version is sent in the form of snmpV1, snmpV2c, or snmpV3. This command uses a target port of 162 by default.	Global Configuration Mode
<code>no snmp trap <notify-name></code>	Deletes the specified SNMP trap.	Global Configuration Mode
<code>snmp system information <contact location name> <information-string></code>	Configure the system information. The configurable system information includes contact, location, and name.	Global Configuration Mode
<code>no snmp system information <contact location name></code>	Deletes certain system configuration information.	Global Configuration Mode
<code>show snmp community</code>	Displays all current community names and the corresponding read and write permission information.	Normal Mode / Privileged Mode
<code>show snmp trap</code>	Displays all current trap names and the corresponding destination IP addresses and version information sent by the trap.	Normal Mode / Privileged Mode
<code>show snmp system information</code>	Displays system information for SNMP settings.	Normal Mode / Privileged Mode

22.3 SNMP Configuration Example

22.3.1 Configuration

Configure a community name named private with read and write permissions.

Configure an SNMP trap named test and send the destination IP as 10.10.10.10. The SNMP version used is 1.

The contact of the configuration system is: E-mail: networks@acb.com.

The location of the configuration system is Shenzhen.

The name of the configuration system is switch.

The switch is configured as follows:

```
Switch#config t
Switch(config)#snmp community private rw
Switch(config)#snmp system information contact
```

E-mail:networks@abc.com

```
Switch(config)#snmp system information location Shenzhen
Switch(config)#snmp system information name switch
```

Chapter 23 RMON Configuration

This chapter includes the following contents:

- RMON Introduction
- RMON Configuration
- RMON Configuration Example

23.1 RMON Introduction

RMON (Remote Monitoring) is a standard monitoring specification used mainly to monitor the data traffic in a network segment or even the whole network. It is one of the most widely used network management standards at present. The RMON specification is an extension of the SNMP MIB, so it is also a MIB, and it is the most important enhancement to the MIB II standard. RMON enables SNMP to monitor remote devices more effectively and proactively.

The RMON monitoring system consists of two parts: a probe (agent or monitor) and a management station. RMON agents store network information in RMON MIBs, which are implanted directly in network devices such as routers and switches. The management station uses SNMP to obtain RMON data information.

The device supports the four most commonly used groups in RMON: (1) Statistics: Provides statistics for each interface. Most of these objects are counters that record the information the monitor collects from the interface.

(2) History group (history): saves the data sampled from the specified interface at a fixed time interval.

(3) Alarm group (alarm): samples the specified data of all interfaces at a fixed time interval, compares it with the set threshold, and triggers the corresponding event when the condition is met.

(4) Event group (event): sets the event; you can choose to log it or issue a Trap.

23.2 RMON Configuration

The RMON commands include the configuration of four groups, View Configuration, and View Data:

Command	Description	CLI Mode
<code>rmon statistics <1-100> (owner WORD)</code>	Enables the configuration of a statistics group with the specified ordinal number for the port. This is an interactive command. Enter the serial number and owner as prompted, where the owner is optional (the same below). The sequence number is the number of the statistics group configuration, and the value range is 1 to 100.	Port Configuration Mode
<code>no rmon statistics <1-100></code>	Cancels the statistics group configuration for the specified sequence number.	Port Configuration Mode

Command	Description	CLI Mode
<code>rmon history <1-100> buckets <1-100> interval <1-3600> (owner WORD)</code>	Configure the history group parameters with the specified sequence number for this port. This is an interactive command. Enter the sequence number, the number of requested buckets, the time interval, and the owner as prompted. SN is the number of the history group configuration, and the value range is 1 to 100. Request Bucket Quantity is the maximum number of saved data, and the value range is 1 to 100. Sampling Interval is in seconds, and the value range is 1 to 3600.	Port Configuration Mode
<code>no rmon history <1-100></code>	Cancels the history group configuration for the specified sequence number.	Port Configuration Mode
<code>rmon alarm <1-60> WORD <1-3600> (absolute delta) rising-threshold <1-2147483647> <1-60> falling-threshold <1-2147483647> <1-60> (owner WORD)</code>	Configure the alarm group parameters for the specified sequence number. This is an interactive command. Enter the serial number, monitoring object, time interval, comparison mode, upper limit threshold, upper limit event serial number, lower limit threshold, lower limit event serial number, and owner as prompted. SN is the number configured for the alarm group, and the value range is 1 to 60. The monitoring object is the OID of a MIB node. The sampling interval is in seconds, and the value range is 1 to 3600. The comparison mode can be absolute or delta, which represent the absolute value (the value of each sampling) and the relative value (the increment of each sampling relative to the previous sampling). The value range of the upper and lower threshold values is 1 to 2147483647. The event must be configured in advance, and the value range of the number is 1 to 60.	Global Configuration Mode
<code>no rmon alarm <1-60></code>	Cancels the alarm group configuration for the specified sequence number.	Global Configuration Mode
<code>rmon event <1-60> (log log-trap WORD none trap WORD) (description WORD) (owner WORD)</code>	Configure the event group parameters for the specified sequence number. This is an interactive command. Enter the sequence number, event type, community name, description, and owner as prompted. SN is the number of the event group configuration, and the value range is 1 to 60. The event type can be log, log-trap, none, or trap. When log-trap or trap is selected, the community name must also be specified (the community name configuration is ignored in this device).	Global Configuration Mode
<code>no rmon event <1-60></code>	Cancels the event group configuration for the specified sequence number.	Global Configuration Mode
<code>show rmon (statistics history-control alarm event) config</code>	View RMON configuration information. This is an interactive command. View objects based on the input prompts.	Global Configuration Mode
<code>show rmon statistics-data interface IFNAME </code>	To view the RMON statistics group data, you must enter the interface name.	Global Configuration Mode
<code>show rmon history-data interface IFNAME </code>	To view RMON history group data, you must enter an interface name.	Global Configuration Mode

23.3 RMON Configuration Example

Enable statistics group configuration on port ge1/1, sequence number 10, owner tereco.

Enable the collection of history group data on port ge1/8. The serial number is 2. At most 80 pieces of data can be saved. The sampling interval is 1 minute. There is no owner.

The event with configuration sequence number 1 is logged and has no owner. Configure the event with sequence number 3 to send a Trap. The community name is public and there is no owner.

Enable the alarm group numbered 5 to monitor the number of bytes received by each port. When the number of bytes per half minute is greater than 1000, a Trap alarm is sent; when it is less than 10, a log is recorded. No owner.

The switch is configured as follows:

```
Switch#configure terminal
Switch(config)#interface ge1/1
Switch(config-ge1/1)#rmon statistics 10 owner tereco
Switch(config-ge1/1)#exit
Switch(config)#interface ge1/8
Switch(config-ge1/8)#rmon history 2 buckets 80 interval 60
Switch(config-ge1/8)#exit
Switch(config)#rmon event 1 log
Switch(config)#rmon event 3 trap public
Switch(config)#rmon alarm 5 1.3.6.1.2.1.2.2.1.10 30 delta rising-threshold 1000 3 falling-
threshold 10 1
```

Chapter 24 Cluster Configuration

Switches provide a cluster management feature that lets a single device manage a group of network devices. This chapter describes how to configure cluster management, including the following:

- Introduction to Cluster Management
- Configure the management device
- Configure Member Device
- Cluster Management Display and Maintenance
- Typical Configuration Example of Cluster Management

24.1 Introduction to Cluster Management

24.1.1 Cluster Definition

A cluster is a collection of network devices that can be managed as a single device.

The purpose of cluster management is to solve the problem of centralized management of a large number of scattered network devices.

Advantages of a cluster: it saves public network IP addresses and simplifies configuration and management tasks. The network manager only needs to configure the public network IP address on one switch in the cluster to manage and maintain the other switches in the cluster.

The switch that is configured with the public network IP address and performs the management function is the command switch; the other managed switches are the member switches. The command switch and the member switches form a "cluster".

The cluster configures and manages the switches in the cluster through the following three protocols.

- NDP (Neighbor Discovery Protocol)
- NTDP (Neighbor Topology Discovery Protocol)
- Cluster (Cluster Management Protocol)

The working process of the cluster includes topology collection and the establishment and maintenance of the cluster. The topology collection process and the cluster maintenance process are relatively independent. The topology collection process starts before the cluster is established. The working principle is as follows:

- All devices obtain information about neighboring devices through NDP, including the software version, host name, MAC address and port name of each neighbor.
- Through NTDP, the management device collects device information within the hop count range you specify, along with the connection information of each device. It then determines the candidate devices of the cluster from the collected topology information.
- Based on the candidate device information collected by NTDP, the management device adds candidate devices to the cluster and removes member devices from it.

All cluster messages are Layer 2 Ethernet messages. For the specific format and interaction process, see the national standard YDT 1692-2007 Technical Requirements for Cluster Management of Ethernet Switches.

24.1.2 Cluster Role

Each device in the cluster takes a different role according to its position and function. You can specify the roles through configuration. The roles are as follows:

1) Command switch:

The only switch in a cluster that can configure and manage the entire cluster. It is also the only switch in the cluster with a public IP address.

- The command switch creates the cluster;
- The command switch discovers and determines a candidate switch by collecting NDP (Neighbor Discovery Protocol) and NTDP (Neighbor Topology Discovery Protocol) information;
- The command switch controls the maintenance of the cluster. It can add a candidate switch to the cluster or delete a member switch from the cluster;
- After the cluster is established, the command switch provides a management channel to the cluster.

2) Member switch:

The managed switch in the cluster.

Member switches are candidates before they join the cluster.

A member switch does not have a public network IP address;

Member switches are managed through the command switch agent.

3) Candidate switch:

A switch that is able to join a cluster but has not yet joined one.

A switch must be a candidate before it can become a member.

4) Independent switch:

A switch that is not cluster-capable.

The roles can be switched according to certain rules:

- When you create a cluster on a candidate device, you designate the current candidate device as the cluster management device. Each cluster must have one and only one management device. Once the management device is designated, it discovers and determines candidate devices by collecting the relevant information. You can then add a candidate device to the cluster through the corresponding configuration.
- After joining the cluster, a candidate device becomes a member device.
- When a member device is deleted from a cluster, it reverts to a candidate device.
- A managed device can revert to a candidate only when the cluster is deleted.

24.1.3 NDP Introduction

NDP obtains information about directly connected neighbor devices, including the connection port, device name and software version. It works as follows:

- A device running NDP periodically sends NDP messages to its neighbors. An NDP message contains the NDP information of the current device (device name, software version, connection port) and the aging time of that information on the receiving device. The device also receives NDP messages sent by neighbor devices, but does not forward them.
- A device running NDP stores and maintains an NDP neighbor information table, creating an entry for each neighbor device. If a neighbor is newly found, that is, an NDP message from that neighbor is received for the first time, a new entry is added to the table. If the NDP information received from the neighbor differs from the old information, the corresponding data item in the table is updated. If the NDP information is the same as the old information, only the aging time is updated. If no NDP information is received from the neighbor before the aging time expires, the corresponding neighbor entry is deleted automatically.

24.1.4 NTDP Introduction

NTDP collects information about each device in a network and about the connections between them. NTDP provides the management device with information about devices that can be added to the cluster, and collects topology information about devices within a specified hop count.

NDP provides the adjacency table information for NTDP. NTDP sends and forwards NTDP topology collection requests according to that adjacency information, and collects the NDP information of each device within a certain network range as well as the connection information between each device and all of its neighbors. Once the information is collected, the management device or the network manager can use it as needed to complete the required functions. When NDP on a member device finds that a neighbor has changed, it notifies the management device of the change through a handshake message. The management device can then start NTDP to collect the specified topology, so that NTDP reflects the change of the network topology in time.

The management device can collect the topology in the network at regular intervals, and you can also start a topology collection once through a manual configuration command. The management device collects the topology information as follows:

- The management device sends the NTDP topology collection request message at regular intervals from the ports on which NTDP is enabled.
- The device that receives the request message immediately sends a topology response message to the management device, and copies the request message to the ports on which NTDP is enabled and sends it to the adjacent devices. The topology response message contains the basic information of the device and the NDP information of all adjacent devices.
- After receiving the request message, each adjacent device performs the same operation, until the topology collection request message has spread to all devices within the specified hop range.
- When the topology collection request message spreads through the network, a large number of network devices receive the request and send topology response messages at the same time. To avoid

network congestion and a busy management device, the following measures can be taken to control the spread speed of the topology collection request message.

- After receiving a topology collection request, the device does not forward the request message immediately. It waits for a certain period of time before forwarding the request message on the ports where NTDP is enabled.
- On the same device, every port with NTDP enabled except the first one waits a certain time after the previous port has sent the topology collection request message before it forwards the message.

24.1.5 Cluster Management and Maintenance

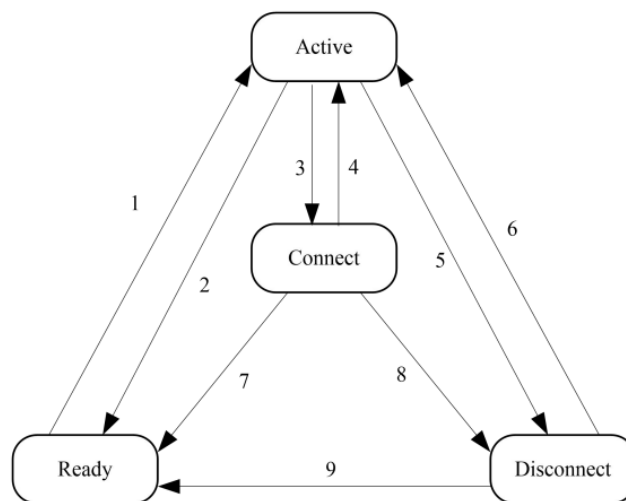
(1) The candidate device joins the cluster

Before establishing a cluster, you must first designate a management device. The management device discovers and determines a candidate device through the NDP and NTDP protocols, and adds the candidate device to the cluster automatically. Alternatively, you can add the candidate device to the cluster through manual configuration.

Before establishing a cluster, you must first designate a management device. The management device discovers and determines a candidate device through the NDP and NTDP protocols, and adds the candidate device to the cluster automatically. Alternatively, you can add the candidate device to the cluster through manual configuration.

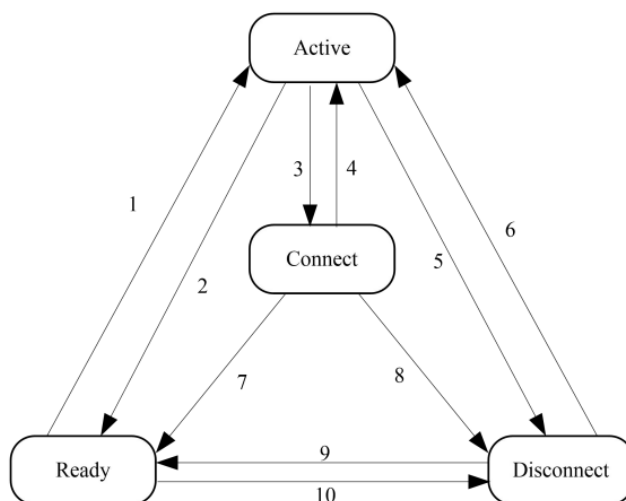
(2) Cluster internal communication

Within the cluster, the management device communicates with the member devices in real time through handshake messages to maintain the connection status between them. The connection status between the management device and the member devices is shown in the following figure.



Command Switch State Transition Diagram

1. Member addition	2. Member deletion	3. No handshake signal received three consecutive times
4. Handshake signal received	5. Recovery request received	6. Interrupt recovery. Re-registration passed
7. Member deletion	8. Status maintained beyond the specified time	9. Member deletion



Member Switch State Transition Diagram

1. Join the cluster	2. Exit the cluster	3. Handshake message not received three consecutive times
4. Handshake message received	5. Join request received	6. Interrupt recovery. Re-registration passed
7. Quit the cluster	8. Status kept beyond the specified time, or a join request message is received	9. Quit the cluster
10. Restore the configuration		

The command switch collects the basic information of the device and identifies it as a candidate switch, which is initially in the Ready state.

Removing a member in any state returns that member switch to the Ready state and identifies it as a candidate switch.

- After a cluster is established successfully and a candidate device joins the cluster to become a member device, the management device saves the state information of the member device locally and marks the

member state as Active. The member device also saves its own state information locally and marks its own state as Active.

- The management device and the member device send handshake messages to each other at regular intervals. After receiving a handshake message from the member device, the management device does not respond and keeps the member device in the Active state. The member device also does not respond and keeps its own state as Active.
- After sending a handshake message to the member device, if the management device does not receive a handshake message from the member device within three times the handshake message sending interval, it changes the locally stored state of the member device from Active to Connect. Similarly, if the member device does not receive a handshake message from the management device within three times the handshake message sending interval after sending its own handshake message, its own state also changes from Active to Connect.
- If the management device receives a handshake message or a management message from a member device in the Connect state within the valid retention time, the state of the member device returns to Active. Otherwise, the state of the member device changes to Disconnect, and the management device considers the member disconnected. If a member device in the Connect state receives a handshake message or management message from the management device within the valid retention time, it changes its own state to Active. Otherwise, it changes to Disconnect.
- When the interrupted communication between the management device and the member device is restored, the member device in the Disconnect state rejoins the cluster. After it joins successfully, the state of the member device on the management device and its local state both return to Active.

If a topology change is found, the member device also sends the change information to the management device through a handshake message.

24.1.6 Management VLAN

The management VLAN limits the scope of cluster management. Configuring the management VLAN provides the following functions:

- The management messages of the cluster (including NDP, NTDP and handshake messages) are confined to the management VLAN and isolated from other messages, which increases security.
- Internal communication between the management device and the member devices takes place through the management VLAN.

Cluster management requires that the ports connecting the management device and the member or candidate devices, including the cascade ports, must allow the management VLAN to pass through. (When a candidate device connects to the management device through another candidate device, the ports connecting the candidate devices are called cascade ports.) Therefore:

- If a port does not allow the management VLAN to pass through, the device connected to that port cannot join the cluster. Before setting up the cluster, make sure that the ports connecting the candidate device and the management device, including the cascade ports, allow the management VLAN to pass through.

- Management VLAN messages are allowed to pass without a tag only when the default VLAN ID of the port connecting the management device and the member or candidate device, and the default VLAN ID of the cascade port, are the management VLAN. Otherwise, management VLAN messages must pass with a tag.

For information about VLANs, see [Configuring VLANs](#).

24.2 Introduction to Cluster Configuration

Before configuring the cluster, define the role and function of each device in the cluster, configure the relevant functions, and plan the communication with the devices in the cluster.

Configuration Task	
Configure the management device	Enable NDP for the system and the port
	Configure NDP parameters
	Enable NTDP for the system and the port
	Configure NTDP parameters
	Configure manual collection of NTDP information
	Enable the cluster function
	Establish the cluster
Configure the member device	Configure interaction between members within the cluster
	Configure cluster member management
	Enable NDP for the system and the port
	Enable NTDP for the system and the port
	Configure manual collection of NTDP information
Configure cluster members to access each other	Enable the cluster function

Notice:

After the cluster is established, turning off NDP or NTDP on the management device and the member devices does not dissolve the cluster, but it does affect the normal operation of the established cluster.

24.3 Configure the Management Device

24.3.1 Enable NDP for the System and the Port

Command	Description	CLI Mode
<code>ndp global enable</code>	Enable the global NDP function. Disabled globally by default.	Configuration Mode
<code>ndp enable</code>	Enable the NDP function of the port. Disabled on all ports by default.	Interface Configuration Mode

Notice:

- NDP must be enabled both globally and on the port for NDP to work properly.
- The NDP feature does not support aggregated ports.
- To prevent the management device from collecting the topology information of a device that does not need to join the cluster and adding it to the cluster, disable NDP on the port connected to that device.

24.3.2 Configure NDP Parameters

Command	Description	CLI Mode
<code>Ndp aging-timer <aging-time></code>	Configure the aging time on the receiving device of the NDP message sent by this device. Default is 180 seconds.	Configuration Mode
<code>ndp hello-timer <hello-time></code>	Configure the interval for sending NDP messages. Default is 60 seconds.	Configuration Mode

Notice:

In general, the aging time of the NDP message on the receiving device must not be less than the NDP sending interval. Otherwise the NDP port neighbor information table becomes unstable.

24.3.3 Enable NTDP for the System and the Interface

Command	Description	CLI Mode
<code>ntdp global enable</code>	Enable the global NTDP function. Disabled globally by default.	Configuration Mode
<code>ntdp enable</code>	Enable the NTDP function of the port. Disabled on all ports by default.	Interface Configuration Mode

Notice:

- NTDP must be enabled both globally and on the port for NTDP to work.
- The NTDP feature does not support aggregated ports.

- To prevent the management device from collecting the topology information of a device that does not need to join the cluster and adding it to the cluster, disable NTDP on the port connected to that device.

24.3.4 Configure NTDP Parameters

Command	Description	CLI Mode
<code>ntdp hop <hop-value></code>	Configure the scope of the topology collection. By default, the farthest device in the collected topology is a maximum of 3 hops from the topology collection device.	Configuration Mode
<code>ntdp timer <interval-time></code>	Configure the interval for scheduled topology collection. Default is 1 minute.	Configuration Mode
<code>ntdp timer hop-delay <time></code>	Configure how long the collected device waits before the first port forwards the topology collection request message. Default is 200 milliseconds.	Configuration Mode
<code>ntdp timer port-delay <time></code>	Configure the port delay time for the current device to forward topology collection requests. Default is 20 milliseconds.	Configuration Mode

24.3.5 Configure Manual Collection of NTDP Information

After the cluster is established, the management device collects the topology information periodically. You can also collect NTDP information manually through configuration at any time, whether or not the cluster is established, and start an NTDP information collection process. This lets you manage and monitor devices in real time more effectively.

Command	Description	CLI Mode
<code>ntdp explore</code>	Collect topology information manually once.	Normal Mode, Privileged Mode

24.3.6 Enable the Cluster Function

Command	Description	CLI Mode
<code>cluster enable</code>	Enable the cluster function. The cluster feature is disabled by default.	Configuration Mode

24.3.7 Set Up a Cluster

The management VLAN limits the scope of cluster management. Configuring the management VLAN provides the following functions:

- The management messages of the cluster (including NDP, NTDP and handshake messages) are confined to the management VLAN and isolated from other messages, which increases security.

- Internal communication between the management device and the member devices takes place through the management VLAN.

Command	Description	CLI Mode
<code>cluster management-vlan <vlan-id></code>	Specify the management VLAN. The default management VLAN is VLAN1.	Configuration Mode

Notice:

If the current device is in the cluster, you cannot modify the management VLAN.

Not in a cluster:

- 1) Check whether the VLAN exists. If it does not, the operation fails directly; otherwise, proceed to the next step.
- 2) Check all the interfaces. If the VLAN where an interface is located and the management VLAN are not the same VLAN, the global NDP and NTDP switches that are turned on are turned off and cleared, and then turned on again.
- 3) Find the Layer 3 interface of the VLAN to be configured. If it is not found, create a new Layer 3 interface for that VLAN. If the creation fails, the management VLAN is still configured successfully. You can use NDP and NTDP, but you cannot join the cluster.
- 4) Set the MAC of the current Layer 3 interface as dev_ID. If the VLAN is set successfully but the new Layer 3 interface fails, use the MAC of vlan1 as dev_ID.

If the management VLAN has been configured but you delete that VLAN directly from the VLAN database, the management VLAN is automatically set to vlan1, and the global NDP, NTDP and cluster switches that were turned on are turned off and cleared accordingly.

Before establishing the cluster, you must first set the private IP address range used by the member devices in the cluster. When a candidate device is added, the management device dynamically allocates a private IP address that can be used within the cluster and sends it to the candidate device for communication within the cluster. This is how the management device manages and maintains the member devices.

Command	Description	CLI Mode
<code>cluster ip-pool <IP/MASK></code>	On the device to be set as the management device, configure the private IP address range used by the member devices in the cluster.	Configuration Mode

Notice:

- The IP address of the VLAN interface of the management device and the member devices and the cluster address pool cannot be configured in the same network segment. Otherwise the cluster does not work properly.
- It can be configured only when the device is not in a cluster.

- The management VLAN is used to search for a corresponding Layer 3 port. If no such port exists, the operation fails directly. (The device cannot be used as a cluster command switch.) If there is a Layer 3 interface, the base address of the IP-Pool is configured on that interface. If this configuration fails, the IP-Pool configuration also fails.

By default, the device is not a management device. The cluster is established as follows:

Command	Description	CLI Mode
<code>cluster build <name></code>	Create a cluster manually, configure the current device as the management device, and assign a cluster name.	Configuration Mode
<code>cluster auto-build <name></code>	Establish the cluster automatically. The auto-clustering feature automatically joins all candidate devices discovered within a specified hop count to the cluster created.	Configuration Mode
<code>cluster delete <name></code>	Delete the cluster.	Configuration Mode
<code>cluster stop auto-add member</code>	In automatic cluster configuration, stop joining member switches automatically. This stops only new devices from joining; devices already in the cluster remain in it.	Configuration Mode

Notice:

- You can specify the management VLAN only before establishing the cluster. After the device has joined the cluster, you cannot modify the management VLAN. To change the management VLAN after the cluster is established, delete the cluster on the management device, designate the management VLAN again, and then re-establish the cluster.
- For security reasons, do not configure the management VLAN as the default VLAN ID for the port where the management device connects to the member device, or for the cascade port.
- Management VLAN messages are allowed to pass without a tag only when the default VLAN IDs of the ports connecting the management device and the member device and of all cascade ports are the management VLAN. Otherwise, those ports and all cascade ports must be configured to allow management VLAN messages to pass with a tag. For the specific configuration, see VLAN.
- The private IP address range of the member devices in the cluster can be configured only before the cluster is established, and only on the management device. If the cluster is already established, the system does not allow the IP address range to be modified.

24.3.8 Configure Intra-Cluster Member Interaction

Within the cluster, the management device communicates with the member devices in real time through handshake messages to maintain the connection status between them. You can configure the handshake message sending interval and the effective retention time of the device on the management device. The configuration takes effect for all member devices in the cluster at the same time.

Command	Description	CLI Mode
<code>cluster timer <interval-time></code>	Configure the interval for sending handshake messages. Default is 10 seconds.	Configuration Mode
<code>cluster holdtime <hold-time></code>	Configure the effective retention time for the device. Default is 60 seconds.	Configuration Mode

24.3.9 Configure Cluster Member Management

On the management device, you can manually add a specified candidate device to the cluster or manually delete a specified member device from the cluster. Cluster member add and delete operations must be performed on the management device; otherwise an error message is returned.

Command	Description	CLI Mode
<code>cluster add member mac-address <mac-address></code>	Add the candidate device to the cluster.	Configuration Mode
<code>cluster delete member mac-address <mac-address></code>	Remove the member device from the cluster.	Configuration Mode

24.4 Configure the Member Device

24.4.1 Enable NDP for the System and the Port

See Enable NDP for the System and the Port.

24.4.2 Enable NTDP for the System and the Port

See Enable NTDP for the System and the Port.

24.4.3 Configure Manual Collection of NTDP Information

See Configure Manual Collection of NTDP Information.

24.4.4 Enable the Cluster Function

See Enable the Cluster Function.

24.5 Configure Access to Cluster Members

After NDP, NTDP and the cluster function are configured correctly, you can configure, manage and monitor the member devices in the cluster from the management device. Switch to the operation interface of a specified member device on the management device to perform configuration management on that member device.

Command	Description	CLI Mode
<code>cluster switch-to member <member-number></code>	Switch from the management device operation interface to the member device operation interface.	Normal Mode, Privileged Mode

Notice:

Switching between the cluster management device and a member device uses a Telnet connection. Note the following during the switchover:

- Before performing the switchover, the peer device must run the "telnet server enable" command to enable Telnet, or the switchover fails.
- When you switch from the management device to a member device, an error message is displayed if member number n does not exist.
- If the Telnet user capacity of the device you are trying to log in to is full, the switchover fails.

24.6 Cluster Management Display and Maintenance

Command	Description	CLI Mode
<code>show ndp[interface <ifname>]</code>	Display NDP configuration information.	Normal Mode, Privileged Mode
<code>reset ndp statistics [interface <ifname>]</code>	Clear NDP statistics.	Configuration View
<code>show ntdp</code>	Display system NTDP information.	Normal Mode, Privileged Mode
<code>show ntdp device-list</code>	Display device information collected by NTDP.	Normal Mode, Privileged Mode
<code>show ntdp single-device mac-address <mac-address></code>	Display NTDP details for the specified device.	Normal Mode, Privileged Mode
<code>show cluster</code>	Display the status and statistics of the cluster the device belongs to.	Normal Mode, Privileged Mode
<code>show cluster topology</code>	Display cluster topology information.	Normal Mode, Privileged Mode
<code>show cluster candidates [mac-address <mac-address>]</code>	Display candidate device information.	Normal Mode, Privileged Mode
<code>show cluster members [<member-number>]</code>	Display cluster member information.	Normal Mode, Privileged Mode

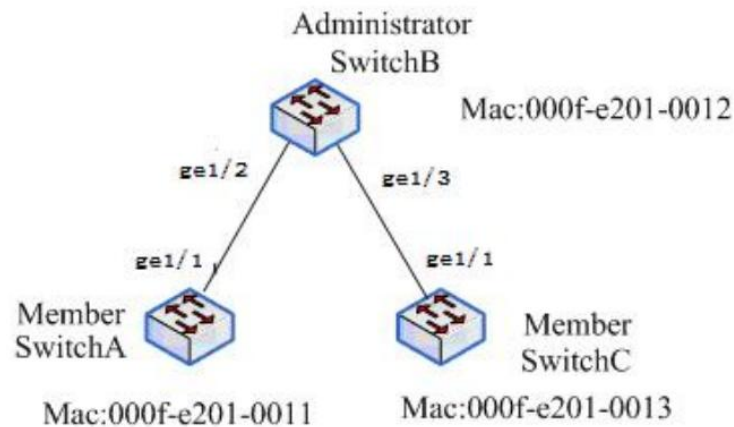
24.7 Cluster Management Configuration Example

(1) Networking requirements

A cluster of three switches, A, B and C, has a management VLAN of VLAN 10. Switch B is the Administrator; Switch A and Switch C are the Members.

The base IP address of the entire cluster address pool is 10.0.0.1, which supports 8 devices.

(2) Networking diagram:



(3) Configuration steps:

Configure member device SwitchA

Configure the management VLAN.

[SwitchA] cluster management-vlan 10

[SwitchA] interface ge1/1

[SwitchA-ge1/1] switch access vlan 10

Enable the global NDP function and the NDP function on port ge1/1.

[SwitchA] ndp global enable

[SwitchA] interface ge1/1

[SwitchA-ge1/1] ndp enable

Enable the global NTDP function and the NTDP function on port ge1/1.

[SwitchA] ntdp global enable

[SwitchA] interface ge1/1

[SwitchA-ge1/1] ntdp enable

Enable the cluster function.

[SwitchA] cluster enable

Configure member device SwitchC

The member devices have the same configuration, so the configuration of Switch C is similar to that of Switch A and is omitted here.

Configure member device SwitchB

Configure the management VLAN.

```
[SwitchB] cluster management-vlan 10
```

```
[SwitchB] interface ge1/2
```

```
[SwitchB-ge1/2] switch access vlan 10
```

```
[SwitchB] interface ge1/3
```

```
[SwitchB-ge1/3] switch access vlan 10
```

Enable the global NDP and NTDP functions, and enable the NDP and NTDP functions on ports ge1/2 and ge1/3.

```
[SwitchB] ndp global enable
```

```
[SwitchB] ntdp global enable
```

```
[SwitchB] interface ge1/1
```

```
[SwitchB-ge1/2] ndp enable
```

```
[SwitchB-ge1/2] ntdp enable
```

```
[SwitchB] interface ge1/3
```

```
[SwitchB-ge1/3] ndp enable
```

```
[SwitchB-ge1/3] ntdp enable
```

Configure the aging time on the receiving device of the NDP message sent by this device to 200 seconds.

```
[SwitchB] ndp timer aging 200
```

Configure the interval for sending NDP messages to 70 seconds.

```
[SwitchB] ndp timer hello 70
```

Configure the maximum number of hops for topology collection to 2 hops.

```
[SwitchB] ntdp hop 2
```

Configure the delay before the first port of the collected device forwards the topology collection request message to 150ms.

```
[SwitchB] ntdp timer hop-delay 150
```

Configure the delay for forwarding the topology collection request message on the other ports of the collected device to 15ms.

```
[SwitchB] ntdp timer port-delay 15
```

Configure the topology collection interval to 3 minutes.

[SwitchB] ntdp timer 3

Enable the cluster function.

[SwitchB] cluster enable

Configure the private IP address range of the member devices as 10.0.0.1 ~ 10.0.0.9.

[SwitchB] cluster ip-pool 10.0.0.1 8

Configure the current device as the management device and establish a cluster named ABC. Members join the cluster automatically.

[SwitchB] cluster autobuild abc

When all the switches you want have joined, you can turn off the automatic cluster joining function.

[SwitchB] cluster stop auto-add member

Chapter 25 SNTP Configuration

This chapter includes the following contents:

- Introduction
- Configuring SNTP
- Show SNTP

25.1 Introduction to SNTP

The Internet generally uses one communication protocol to achieve network time synchronization, namely NTP (Network Time Protocol). The other protocol is a simplified version of NTP: SNTP (Simple Network Time Protocol).

NTP runs across a variety of platforms and operating systems. Its algorithms are very sophisticated, so it is almost unaffected by network latency and jitter and can provide 1-50ms accuracy. NTP also provides an authentication mechanism with a high level of security. However, the NTP algorithm is complex and places high demands on the system.

SNTP (Simple Network Time Protocol) is a simplified version of NTP. It uses a simple algorithm to calculate the time, so its performance is high. Its accuracy is generally about 1 second, which meets the needs of most situations.

Because the SNTP message is completely consistent with the NTP message, the SNTP Client implemented by the switch is fully compatible with an NTP Server.

25.2 Configuring SNTP

25.2.1 Default SNTP Settings

Item	Default Value
SNTP Status	Disable, the SNTP service is off
NTP Server	None
SNTP time synchronization interval	1800 seconds
Local Time Zone	+8, that is, East 8th District

Turn SNTP on and off

Configure it as follows:

```
Switch# configure terminal
```

Enter global configuration mode

```
Switch(config)# sntp enable
```

Turn SNTP on

```
Switch(config)# sntp disable
```

Turn SNTP off

25.2.2 Configure the SNTP Server Address

Because the SNTP message is completely consistent with the NTP message, the SNTP Client is fully compatible with an NTP Server. There are many NTP Servers on the network, so you can select one with less network latency as the NTP Server on the switch.

You can obtain a specific NTP server address by logging in to <http://www.time.edu.cn/> or <http://www.ntp.org/>. For example, 192.43.244.18 (time.nist.gov).

The switch can be configured with three server addresses at most. The switch uses the first Server address to synchronize the time. If synchronization fails, it uses the second Server address, and so on.

Add a Server address as follows:

```
Switch# configure terminal
```

Enter global configuration mode

```
Switch(config)# sntp server 210.72.145.44
```

Add the SNTP server IP. If the switch already has three Server addresses, the addition fails; delete an address before adding a new one. Delete a Server address as follows:

```
Switch(config)# no sntp server
```

Delete all Server addresses

```
Switch(config)# no sntp server 210.72.145.44
```

Delete one Server address

25.2.3 Configure the SNTP Clock Synchronization Interval

The SNTP Client must synchronize with the NTP Server periodically to keep its clock synchronized.

Configure it as follows:

```
Switch# configure terminal
Switch(config)# sntp interval 60
```

Set the interval of the timing synchronization clock, in seconds, from 60 seconds to 65535 seconds. The default value is 1800 seconds; here it is set to 60 seconds.

```
Switch(config)# no sntp interval
```

Restore the interval of the timing synchronization clock to the default of 1800 seconds

25.2.4 Configure the Local Time Zone

The time obtained through the SNTP protocol is Greenwich Mean Time (GMT). To obtain the local time, you must set the local time zone so that the standard time is adjusted. By default, the local time zone of the switch is East Zone 8, which is the time zone of China.

Configure it as follows:

```
Switch# configure terminal
Switch(config)# sntp time-zone -8
```

Set the local time zone to West Zone 8

```
Switch(config)# no sntp time-zone
```

Restore the local time zone to East 8

25.3 Information Display of SNTP

Configure it as follows:

```
Switch# show sntp
Switch# show running-config
```

Chapter 26 Configuring PTP

This chapter includes the following contents:

- Introducing PTP
- Configuring PTP

26.1 Introducing PTP

PTP: Precision Time Protocol. As network technology has grown, and as Ethernet in particular is used more and more widely in measurement and control systems, the computer and network industry has also worked to solve the insufficient timing synchronization capability of Ethernet, in order to reduce the use of other technologies such as IRIG-B. A software-based Network Time Protocol (NTP) was therefore developed to improve the timing synchronization capability among network devices.

PTP is suitable for distributed network communication systems that support multicast messaging, such as EtherNet. Unicast messages are also supported. The protocol supports multiple transport protocols, such as UDP/IPv6, Layer-2 EtherNet and DeviceNet. It uses short frame transmission, has few data frames, uses a simple algorithm, consumes few network resources and has low computing performance requirements, so it suits low-end equipment.

The IEEE 1588 standard has two versions, Version 1 and Version 2. Version 1 was released at the end of 2002. The latest Version 2 is IEEE1588-2008, which provides all the functions of Version 1 and improves on its shortcomings.

Clock synchronization implemented by the PTP protocol has two main steps:

1. Establish a synchronous system, that is, establish a master-slave synchronous system within the whole synchronous system through an optimal master clock algorithm;
2. During the synchronization process, calculate and synchronize the local clock by exchanging PTP protocol messages.

26.2 Configuring PTP

26.2.1 PTP Settings

Item	Default Value
PTP status	Disable, the PTP service is off

Turn PTP on and off

Configure it as follows:

```
Switch# configure terminal
```

Enter global configuration mode

```
Switch(config)#ptp tc enable
```

Turn on SNTP

```
Switch(config)#no ptp tc enable
```

Turn off PTP

Chapter 27 Configuring IGMP

This chapter covers the following topics:

- IGMP introduction
- IGMP configuration
- IGMP configuration example

27.1 IGMP Introduction

IGMP (Internet Group Management Protocol) is part of multicast technology. It runs between a host and its directly connected multicast router. On one side, the host reports to the local router the multicast group it wants to join or leave. On the other side, the multicast router can periodically query whether the local host is actively receiving a specific group. IGMP currently has three versions: version 1 (rfc1112), version 2 (rfc2236) and version 3 (rfc3376). IGMP is based on a query and response mode. The multicast router uses periodic query messages to ask whether a host in the subnet wants to receive a specific group. The host uses a report message to tell the local router which multicast group it wants to join or leave, that is, the membership report. When a host wants to join a specific group, it does not have to wait for the periodic query message; it can send the group membership report to the router immediately. When the router queries the membership of the hosts in the subnet, each host uses a suppression mechanism and delays its report message by a random value between 1 and 10 seconds. If the host receives the same group report message from another host during that delay, it suppresses its own group member report, which reduces IGMP traffic in the subnet. When there are multiple routers in the subnet, their IP addresses are compared to select one as the querier, so that the subnet does not carry too many IGMP query messages. Version 1 supports only query and report. When a group member leaves, the multicast router stops forwarding the multicast data stream only after the group information times out. The querier is elected by other protocols. Version 2 adds a leave message, which supports the immediate departure of group members. When the last group member leaves the group, it sends a leave message to the router, and the router sends a specific group query message to determine whether other receiving hosts remain in the subnet. If the corresponding group membership message is still not received after the configured number of retransmissions, the group becomes invalid and forwarding of the multicast data stream stops. Version 2 effectively reduces the leave delay through this leave mechanism. Version 2 also provides a querier election mechanism. When other routers in the subnet receive a general query message, they compare their own IP address and select the one with the smallest IP address as the querier, and start a timer. The timer resets each time a query message is received. When the timer expires, a new election begins. To adjust the response time of the host suppression mechanism, version 2 adds a maximum response time field to the message. The querier can control the suppression response time of the host by setting this value in the query message. When a host receives the query message, it uses the maximum response time to generate a local delay and starts a timer. If it receives the same group membership message from another host within that delay, its own report message is suppressed. Version 3 supports source-specific multicast. Query messages are divided into general query, specific group query and specific group specific source query. The report message expands the original group address field into group record entries. Each group record contains the group address and its corresponding source address list. Whether a specific source is joined is distinguished by the

group record type. Icmpv3 works with pim-ssm to implement source-specific multicast. The SPT is established directly from the source to the receiver, without the convergence point RP.

27.2 IGMP Configuration

Most IGMP functions relate to interfaces. Configure the corresponding parameters and timer values for each interface in interface mode.

IGMP configuration includes:

- Start the IGMP function of the interface
- Configure the group filter access control list for the interface
- Configure the access control list for leave-group filtering on the interface
- Configure the number of specific group queries for the interface
- Configure the specific group query interval of the interface
- Configure the non querier timer time of the interface
- Configure the query timer interval of the interface
- Configure the maximum response time of the interface
- Configure interface parameters
- Configure the protocol version of the interface

27.2.1 Start IGMP Function of Interface

Mode: Interface Configuration Mode

Command: IP IGMP Open IGMP protocol on the interface

Command: no IP IGMP Turn off IGMP protocol on the interface

Default: the IGMP protocol is not open;

Use this command to let the interface send and receive IGMP protocol messages. If the IP multicast routing command has enabled multicast support in global configuration mode, then starting the PIM SM protocol with the IP PIM sparse mode command in interface configuration mode automatically starts the IGMP protocol function on that interface. Similarly, in interface configuration mode, the no IP PIM sparse mode command automatically turns off the IGMP protocol function on that interface.

27.2.2 Configure the Group Filter Access Control List for an Interface

Mode: Interface Configuration Mode

Command: IP IGMP access group {< ACL ID > | < ACL name >}

Configure the group filter access control list for the interface

Command: no IP IGMP access group

Delete the group filter access control list for this interface

Parameter: < ACL ID > is the standard access control list number, an integer between 1 and 99. < ACL name > is the standard access control list name. Use an ACL to filter the multicast address table learned by the interface. For ACL configuration details, see the Command Reference Manual.

Default: no group filtering ACL is configured

27.2.3 Configure the Access Control List for Interface Leave Filtering

Mode: Interface Configuration Mode

Command: IP IGMP immediate leave group list {< acl-id1 > | < acl-id2 > | < ACL name >}

Configure the access control list filtered by the interface leaving the group. Command: no IP IGMP immediate leave

Delete the leave-group filter access control list of this interface

Parameter: < ACL Id1 > is the standard access control list number, an integer between 1 and 99. < Acl-id2 > is the number in the extended range of the standard access control list, an integer between 1300 and 1999. < ACL name > is the standard access control list name. Use the ACL to filter the multicast addresses that must be filtered when the interface receives version 2 and version 3 leave messages. For ACL configuration details, see the Command Reference Manual.

Default: no leave group filtering ACL is configured

27.2.4 Configure the Number of Group-Specific Queries for an Interface

Mode: Interface Configuration Mode

Command: IP IGMP last member query count < count >

Configure the number of interface specific group queries

Command: no IP IGMP last member query count

Restore the number of specific group queries on the interface to the default value

Parameter: < count > is the number of specific group queries or specific source queries to send when a leave message is received, an integer between 2-7. If no group report message is received during this period, the group is considered to have no receiving host.

Default: 2

27.2.5 Configure the Group-Specific Query Interval for an Interface

Mode: Interface Configuration Mode

Command: IP IGMP last member query interval < interval >

Configure the specific group query interval of the interface

Command: no IP IGMP last member query interval

Restore the specific group query interval of the interface to the default value

Parameter: < interval > is the interval between a specific group query or a specific group source query, an integer between 1000-25500, in milliseconds. When a leave message is received, a specific group query or a specific source query for that group must be sent several times. This command configures the sending interval of those query messages. If no group report message is received after the queries are sent, the group, or the source of the group, is considered to have no receiving host.

Default: 1 seconds

27.2.6 Configure the Non Querier Timer Time of the Interface

Mode: Interface Configuration Mode

Command: IP IGMP query timeout < time >

Configure the non querier timer time of the interface

Command: no IP IGMP query timeout

Restore the non querier timer time of the interface to the default value

Parameter: < time > is the expiry time of the non querier timer, an integer between 60-300. When there are multiple switches on the subnet, one must be selected as the querier to send query messages. During network initialization, all switches send query messages by default. When a switch receives query messages from other switches, the IP addresses are compared and the one with the smaller IP address is selected as the querier. A switch that is not the querier starts a timer, which is reset each time a query message is received. If the timer expires, the querier is elected again.

Default: 255 seconds

27.2.7 Configure the Query Timer Interval of the Interface

Mode: Interface Configuration Mode

Command: IP IGMP query interval < interval >

Configure the query timer interval of the interface

Command: no IP IGMP query interval

Restore the query timer interval of the interface to the default value

Parameter: < interval > is the interval of the query timer, an integer between 1-18000. When the interface starts the IGMP protocol, it starts a query timer to send query messages to the subnet regularly. If there are multiple switches, a querier must be elected. The querier uses the timer to send query messages periodically. A non querier closes the timer and starts a timeout timer. When the timeout timer expires, the querier is elected again.

Default: 125 seconds

27.2.8 Configure the Maximum Response Time for an Interface

Mode: Interface Configuration Mode

Command: IP IGMP query Max response time < time >

Configure the maximum response time carried by the interface when it sends a query message

Command: no IP IGMP query Max response time

Restore the maximum response time carried by the interface when it sends a query message to the default value

Parameter: < time > is the maximum response time value the interface carries when it sends a version 2 or version 3 query message, an integer between 1-240, in seconds. Note that the unit of the maximum response time inside the query message is 0.1 seconds. IGMP uses the query response mode to obtain information about the receiving hosts of the corresponding group on the subnet. When a switch sends a query message, multiple hosts on the network respond with group report messages. To avoid congestion from many IGMP messages, the query message carries a maximum response time value. A host that receives the query message generates a random value based on the maximum response time and delays its response message by that value. If the host receives an identical group report message from another host during this delay, it suppresses the group report message it was going to send, which effectively avoids simultaneous IGMP group report conflicts. Note that the maximum response time range of version 3 is larger than that of version 2, and the maximum response time range of version 2 is less than 128.

Default: 10 seconds

27.2.9 Configure Interface Parameters

Mode: Interface Configuration Mode

Command: IP IGMP robustness variable < value >

Configure interface parameters

Command: no IP IGMP robustness variable

Restore the vitality parameter of the interface to the default value

Parameter: < value > is the vitality parameter of the interface, an integer between 2 and 7. The vitality parameter describes the robustness of the protocol, that is, how well the protocol adapts to packet loss when the network environment is bad. The configured vitality parameter value affects the number of retransmissions and the intervals of the relevant timers. The version 3 query message contains a QRV value, which is used to synchronize the vitality parameter configured locally on non queriers.

Default: 2

27.2.10 Configure the Protocol Version of the Interface

Mode: Interface Configuration Mode

Command: IP IGMP version < version >

Configure the protocol version of the interface

Command: no IP IGMP version

Restore the protocol version of the interface to the default value

Parameter: < version > is the IGMP version number used, an integer between 1 and 3.

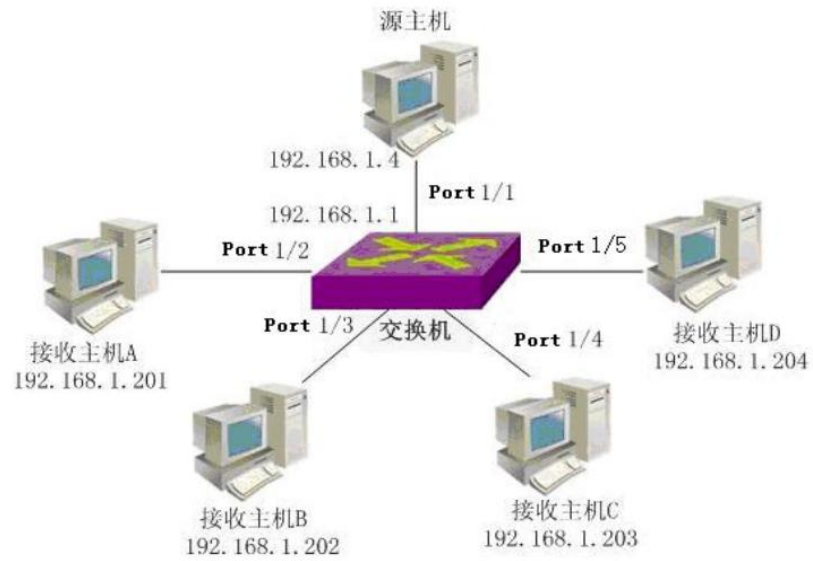
Default: 3

27.3 IGMP Configuration

(1) Configuration

IGMP generally works together with a multicast routing protocol, for example PIM-SM. See the PIM-SM configuration example. Starting the multicast function (IP multicast routing) and the PIM-SM protocol (IP PIM sparse mode) on an interface starts the corresponding IGMP protocol automatically, so you do not have to configure the IP IGMP command separately. When the multicast on demand service flow is received, you can view the corresponding IGMP group information and interface status. When the source host and the receiving host are in the same subnet, the multicast data flow does not have to be forwarded across network segments and Layer 3 multicast routing is unnecessary; you can start the IGMP protocol and IGMP snooping separately to achieve multicast within the subnet. Note that the PIM-SM protocol must be started when writing the Layer 3 multicast hardware table.

```
Switch# configure terminal
Switch(config)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.1/24
Switch(config-vlan2)#ip igmp
Switch(config-vlan2)#ip pim sparse-mode passive
Switch(config-vlan2)#interface ge1/1
Switch(config-ge1/1)#switch access vlan 2
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switch access vlan 2
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#switch access vlan 2
Switch(config-ge1/3)#interface ge1/4
Switch(config-ge1/4)#switch access vlan 2
Switch(config-ge1/4)#interface ge1/5
Switch(config-ge1/5)#switch access vlan 2
```



(2) verification

Use the following command to view IGMP information:

```
show ip igmp group
show ip igmp interface
```

Chapter 28 Configuring PIM-SM

This chapter includes the following contents:

- PIM-SM introduction
- PIM-SM Configuration
- PIM-SM Configuration Example

28.1 PIM-SM Introduction

PIM-SM (Sparse Mode Protocol Independent Multicast - Sparse Mode) is part of multicast technology. It is used for network topology discovery, and it is a protocol that runs between routers. Compared with unicast point-to-point transmission, multicast point-to-multipoint transmission builds a distribution tree so that the source data flow can travel along the tree to the group member hosts that need to receive it. There are two forms of distribution tree. One is source-based: each source has a distribution tree for each group, taking the shortest path from the source to the group members. This active tree is also called the shortest path tree, SPT. The other shares one information distribution point: all sources of a group share a tree, and the information of a source must be registered at this convergence point (RP: rendezvous point) before it can be sent down the distribution tree to the receiving hosts. This shared tree is also known as the RPT. The distribution tree technology of PIM-SM uses a one-way shared tree, that is, the data flow can only travel from the source to the receiver, not the other way round. PIM-SM is designed for the WAN, where the receivers are sparsely distributed in the network, which differs from the flood and prune mode of dense mode. PIM-SM uses the join and prune mode to build a shared tree, that is, the receiving host uses an explicit join (the traditional IP multicast mode, in which the receiving end starts membership). The shared tree is a distribution tree shared by all sources of a group. When different sources need to push data streams, the source registers at the convergence point, and the convergence point then pushes the data to the receiver. The tree from the source to the convergence point is the shortest path tree started by the data. When many sources register at the convergence point to forward data streams, the convergence point becomes a bottleneck in the network. PIM-SM then allows each source and group to switch to its own shortest path tree, bypassing the convergence point. In a small network, one convergence point is enough. In a larger network, multiple RPs must be configured to share the load, but static configuration does not adapt to dynamic changes in the network. PIM-SM therefore provides an RP bootstrap mechanism. Multiple candidate BSRs and candidate RPs are configured in the network. The candidate BSRs elect the BSR. Each candidate RP regularly sends its own group mapping information to the BSR. The BSR collects the RP information, arranges it into bootstrap information, and diffuses it through the network regularly. Each router saves a copy of the bootstrap information. When a group member joins, the router looks up the bootstrap information, finds the RP corresponding to the group, and sends a join message towards that RP. If the bootstrap information contains no RP for the group, the router uses a hash algorithm to map the group to an RP and sends a join message towards that RP. The assertion mechanism and the network sharing mechanism of pim-dr are considered at the same time: the designated router DR applies to the process of joining the distribution tree, and the assertion applies to the process of forwarding the data flow. Because of its careful design, PIM-SM has become the first choice of intra-domain multicast routing protocol. Later, pim-ssm was added so that users can receive services directly from multicast sources, a model whose advantages are more pronounced than traditional PIM-SM.

The concept of multicast security addresses the use of multicast protocols on a Layer 3 switch. Multicast technology can be divided into three parts: multicast source discovery, multicast receiver discovery, and topology discovery. Some multicast protocols run between a host and a router, such as IGMP; others run between routers, such as PIM-SM. Whether it runs between a host and a router or between routers, a multicast protocol uses a Layer 3 interface to describe the input and output direction of a multicast route. On a Layer 3 switch, if a Layer 3 interface is the output interface of a multicast route, every Layer 2 port in the VLAN that corresponds to that interface becomes an output port. When there is only one VOD host in the downstream VLAN, the other output ports in the VLAN are unnecessary. The IGMP snooping protocol therefore runs at Layer 2, where it monitors the IGMP messages received on a port and records them. When a data stream is output from the VLAN, it is sent only to the ports that requested it. That is, the Layer 2 port information must be recorded on the Layer 3 switch. Between routers, join and prune messages are needed to maintain the multicast distribution tree. If the ports that send and receive PIM protocol messages are also monitored, the output interface of the multicast routing table records the corresponding Layer 2 ports as well. A data stream that flows through the multicast tree is then sent only to the ports in the VLAN where a join message arrived, not to all ports in the VLAN. The Layer 2 protocol that monitors PIM protocol messages is called PIM snooping. With IGMP snooping and PIM snooping, a multicast routing protocol that runs on a router also runs well on a Layer 3 switch. The entries in the multicast routing table are shown as follows:

Multicast source address multicast group address input interface output interface list output port list

When an on-demand data stream is present, the branch of the multicast distribution tree from the source to the receiving host follows the Layer 2 ports where the protocol messages once appeared, rather than flooding the VLAN that corresponds to the Layer 3 interface.

IGMP snooping and PIM snooping are dynamic monitoring protocols that record Layer 2 port information from the protocol messages. Currently, switch v3 version 0 provides only the static configuration function that refines multicast routing entries down to Layer 2 ports. This function module is called multicast security. Static configuration commands specify that a multicast data stream is output only to a specific Layer 2 port in the VLAN that corresponds to the output interface. By default, the data flow is output to all Layer 2 ports in the VLAN that corresponds to the output interface.

28.2 PIM-SM Configuration

After the Layer 3 multicast function is started, you must start the PIM-SM protocol on each interface individually. Its commands are in global configuration mode and interface configuration mode.

The configuration of PIM-SM includes:

Start the multicast routing function

Configure the multicast routing table capacity

Configures the TTL value of the multicast interface

Start the interface PIM SM function

Configure passive mode on the interface

Configure the interface priority

Configure the interface Hello message not to contain genid information

Configure the interface Hello timer interval

Configure the hold time of neighbors on the interface

Configure neighbor list filtering on the interface

Configure the source address of the unicast registration message

Configure the limit on the number of registration messages

Configure RP checking on registration

Configure the registration suppression timer value

Configure the registration KAT timer value

Configure registration source address filtering

Configure the checksum of the registration message in Cisco mode

Configure a static RP address

Configure a candidate RP

Configure ignoring of the RP set priority

Configure the c-rp-adv message in Cisco mode

Configure a candidate BSR

Configure the JP timer interval

Configure SPT (shortest path tree) switching

Configures SSM (source specific multicast)

28.2.1 Start Multicast Routing Function

Mode: Global Configuration Mode

Command: IP multicast routing Enables multicast routing

Command: no IP multicast routing Turns off the multicast routing function

Default: the multicast routing function is not started. Use this command to start the multicast routing function and the PIM-SM and IGMP protocols.

28.2.2 Configure Multicast Routing Table Capacity

Mode: Global Configuration Mode

Command: IP multicast route limit < limit > [threshold]

Configures the capacity and alarm threshold of the multicast routing table

Command: no IP multicast route limit

Restores the capacity and alarm threshold of the multicast routing table to their default values

Parameter: < limit > is the capacity of the multicast routing table, an integer between 1-2147483647; [threshold] is the alarm threshold of the multicast routing table, an integer between 1-2147483647, and is optional. When the multicast routing table exceeds the alarm threshold, the switch displays a prompt message.
Default: 2147483647

28.2.3 Configure TTL Value of Multicast Interface

Mode: Interface Configuration Mode

Command: IP multicast TTL threshold < threshold >

Configures the TTL value of the multicast interface

Command: no IP multicast TTL threshold

Restores the multicast interface TTL value to invalid

Parameter: < threshold > is the TTL value of the multicast interface, an integer between 0-255; 256 is an invalid value. This command configures the TTL value of the interface and is used to maintain the output interface list of the multicast routing table. When the configured TTL value is between 0-255, the value is passed to the actual multicast interface. When the value is 256, the actual multicast interface uses its own default TTL value of 1.

Default: 256

28.2.4 Start Interface PIM SM Function

Mode: Interface Configuration Mode

Command: IP PIM spark mode

Starts the interface PIM SM protocol

Command: no IP PIM spark mode

Closes the interface PIM SM protocol

Default: the interface does not start the PIM SM protocol. Use this command to start the PIM SM protocol function on the interface, send and receive protocol messages, and participate in RP election when bootstrapping is configured. When the interface starts PIM SM, the IGMP protocol starts with it. You do not need to configure the IP IGMP command.

28.2.5 Configure Passive Mode of Interface

Mode: Interface Configuration Mode

Command: IP PIM spark mode passive

Configures the interface to passive mode

Command: no IP PIM spark mode passive

Cancels passive mode on the interface and restores it to the active state

Default: the PIM SM protocol is not started. When you start the PIM SM function of the interface with the IP PIM spark mode command, the interface is in the active state and sends and receives protocol messages normally. When passive mode is configured, the interface does not send protocol messages, does not participate in the registration process, does not participate in the RP election process, and only monitors protocol messages quietly.

28.2.6 Configure Interface Priority

Mode: Interface Configuration Mode

Command: IP PIM Dr priority < priority >

Configures the priority of the interface in DR election

Command: no IP PIM Dr priority [priority]

Restores the interface priority to the default

Parameter: < priority > is the priority of the interface in DR election, an integer between 0-4294967294. When there are multiple switches on a shared network, one is elected as the DR through the Hello protocol; it is responsible for sending join and prune messages upstream. In DR election, the priorities are compared first and the higher priority is selected. If the priorities are the same, the switch with the larger IP address is selected as the DR.

Default: 1

28.2.7 Configure the Interface to Omit genid Information from Hello Messages

Mode: Interface Configuration Mode

Command: IP PIM exclude genid

Configures the interface Hello message not to contain the genid field

Command: no IP PIM exclude genid

Restores the default, in which the interface Hello message contains genid

Default: contains the genid field. Type 20 in the Hello message option list is the option value of the generation ID, an unsigned 32-bit value generated randomly when the interface starts. It is used to identify quickly whether a neighbor has restarted, and to reduce the delay of relearning RP set information and join and prune state caused by a router restart. The switch retains the genid value in the Hello message and updates this value after receiving a new Hello message. If the genid changes, the neighbor is considered to have restarted and must be informed of the relevant information in time, instead of waiting for the timer to terminate before learning by itself, which reduces the convergence time when the network changes.

28.2.8 Configure Interface Hello Timer Interval

Mode: Interface Configuration Mode

Command: IP PIM Hello interval < interval

Configures the Hello timer interval of the interface

Command: no IP PIM Hello interval

Restores the interface Hello timer interval to the default value

Parameter: < interval > is the interval at which the interface sends Hello messages, an integer between 1-65535. The Hello timer interval affects the holdtime value, which is 3.5 times the Hello interval. When the holdtime is not configured, or the configured value is less than the Hello interval, it is calculated from the Hello interval.

Default: 30 seconds

28.2.9 Configure the Neighbor Hold Time for an Interface

Mode: Interface Configuration Mode

Command: IP PIM Hello holdtime < value >

Configures the holding time of neighbors that send Hello messages

Command: no IP PIM Hello holdtime

Restores the neighbor holding time to the default value

Parameter: < value > is the survival time of neighbors, an integer between 1-65535. If no Hello message is received within the hold time, the neighbor is considered to no longer exist. The configured value must be greater than the Hello interval. If it is less than the Hello interval, 3.5 times the Hello interval is used as the hold time

Default: 105 seconds (3.5 times the Hello interval)

28.2.10 Configure Neighbor List Filtering of the Interface

Mode: Interface Configuration Mode

Command: IP PIM neighbor filter {< ACL ID > | < ACL name >}

Configures the neighbor list filtering ACL of the interface

Command: no IP PIM neighbor filter {< ACL ID > | < ACL name >}

Cancels neighbor list filtering on the interface

Parameter: < ACL ID > is the number of the standard access control list, an integer between 1 and 99; < ACL name > is the standard access control list name. Use an ACL to filter the neighbor list of the interface. When a neighbor is filtered by the ACL, it is deleted from the neighbor list of the interface. Refer to the Command Reference Manual for ACL configuration details.

Default: no neighbor filtering ACL is configured

28.2.11 Configure the Source Address of Unicast Register Messages

Mode: Global Configuration Mode

Command: IP PIM register source {< IP > | < if name >}

Configures the source address or interface name of the unicast registration message

Command: no PIM register source

Cancels the configured registration source address

Parameter: < IP > is the source address of the unicast registration message, in dotted decimal format; < If name > is the source interface name of the unicast registration message, which is a Layer 3 interface name (for example, vlan2). This is the address used as the source of the registration message.

Default: the source address of the registration message is not configured, and the interface address of the received packet is used.

28.2.12 Configure the Number Limit of Registered Message

Mode: Global Configuration Mode

Command: IP PIM register rate limit < limit >

Configures the threshold for the number of registration messages received within one second

Command: no IP PIM register rate limit

Cancels the check on the number of registration messages

Parameter: < limit > is the maximum number of registration messages that can be received within 1 second, an integer between 1-65535. A registration message encapsulates the data packet. When a source registers, a large data flow hits the first hop router of the source. The data flow of the PIM SM protocol can then be limited to avoid placing a large burden on the router. Because PIM SM only needs the data flow to trigger SPT establishment from the source to the RP, this restriction mechanism can be used. When the data flow is limited, the number of data packets within 1 second is checked. If the arrival interval of two data packets exceeds 1 second, there is no limit.

Default: no limit on the number of registration messages is configured (the threshold value is 0)

28.2.13 Check RP When Configuring Registration

Mode: Global Configuration Mode

Command: IP PIM register RP reachability

Checks whether the RP is reachable when configuring registration

Command: no IP PIM register RP reachability

Does not check whether the RP is reachable when configuring registration

Default: there is no need to check whether the RP is reachable during registration

28.2.14 Configure Registration Suppression Timer Value

Mode: Global Configuration Mode

Command: IP PIM register suppression < value >

Configure the registration suppression timer value

Command: no IP PIM register suppression

Restores the registration suppression timer to the default value

Parameter: < value > is the time value of the registration suppression timer, an integer between 1-65535.

Default: 60 seconds

28.2.15 Configure and Register Kat Timer Time Value

Mode: Global Configuration Mode

Command: IP PIM RP register Kat < value >

Configures the Kat timer time of the (s, g) entries registered and created on the RP Command: no IP PIM RP register Kat

Restores the Kat timer time of the (s, g) entries to the default value

Parameter: < value > is the time value of the Kat timer for the (s, g) entries created by source registration on the RP, an integer between 1-65535.

Default: 185 seconds (3 * register-suppression + register-probe)

28.2.16 Registration Address Configuration

Mode: Global Configuration Mode

Command: IP PIM accept register list {< acl-id1 > | < acl-id2 > | < ACL name >

Configures the source filter list of registration messages

Command: no IP PIM accept register

Cancels the source filter list of registration messages

Parameter: < ACL Id1 > is the number of the extended access control list, an integer between 100 and 199; < Acl-id2 > is the number of the extended part of the extended access control list, an integer between 2000-2699; < ACL name > is the name of the access control list. The access control list can be used to filter the source addresses registered to the RP. If a source address on the RP is filtered by the ACL, the corresponding (s, g) entry is deleted from the multicast routing table. Refer to the Command Reference Manual for ACL configuration details.

Default: no ACL for the registration source is configured

28.2.17 Configure the Register Message Checksum in Cisco Mode

Mode: Global Configuration Mode

Command: IP PIM Cisco register checksum [group list {< ACL Id1 > | < ACL Id2 > | < ACL name >}]

Configures the checksum of the registration message to Cisco mode

Command: no IP PIM Cisco register checksum [group list {< ACL Id1 > | < ACL Id2 > | < ACL name >}]

Cancels the Cisco mode checksum of the registration message

Parameter: < ACL Id1 > is the number of the standard access control list, an integer between 1-99; < Acl-id2 > is the number of the extended part of the standard access control list, an integer between 1300 and 1999; < ACL name > is the name of the standard access control list. The Cisco mode checksum is calculated over the whole registration message, including the data packet part, while the checksum of a general registration message covers only the PIM SM protocol message header and the registration message header. If an access control list is configured, the ACL filters the corresponding group address when the registration message is sent. The Cisco checksum can be used through an ACL, but the general checksum cannot.

Default: the Cisco checksum is not configured

28.2.18 Configure Static RP Address

Mode: Global Configuration Mode

Command: IP PIM RP address < IP > [< acl-id1 > | < acl-id2 > | < ACL name >]

Configures a static RP

Command: no IP PIM RP address < IP > [< acl-id1 > | < acl-id2 > | < ACL name >]

Cancels the static RP configuration

Parameter: < IP > is the address of the static RP, in dotted decimal format; < ACL Id1 > is the number of the standard access control list, an integer between 1 and 99; < Acl-id2 > is the number of the extended part of the standard access control list, an integer between 1300 and 1999; < ACL name > is the name of the standard access control list.

Default: no static RP is configured

28.2.19 Configure Candidate RP

Mode: Global Configuration Mode

Command: IP PIM RP candidate < if name > [group list {< ACL ID > | < ACL name >} | interval < value1 > | priority < Value2 >]

Configures the interface as a C-RP

Command: no IP PIM RP candidate [if name]

Cancels the C-RP attribute of the interface

Parameter: < if name > is the name of the interface configured as C-RP, which is a Layer 3 interface name (for example, vlan2); < ACL ID > is the number of the standard access control list, an integer between 1 and 99; < ACL name > is the name of the standard access control list; < Value1 > is the interval at which the C-RP periodically unicasts a c-rp-adv message to the BSR, an integer between 1-16383; < Value2 > is the priority of the C-RP, an integer between 0 and 255. The priority is included in the bootstrap information. It is flooded in the domain and used to select the RP that corresponds to the group.

Default: value1 defaults to 60 seconds; Value2 defaults to 192.

28.2.20 Configure Ignore RP Set Priority

Mode: Global Configuration Mode

Command: IP PIM ignore RP set priority

Ignores the priority and uses hash algorithm mapping when searching for an RP in the RP set

Command: no IP PIM ignore RP set priority

Uses the priority when searching for an RP in the RP set

Default: the RP is searched for in the RP set in priority order

28.2.21 Configure C-rp-adv Message in Cisco Mode

Mode: Global Configuration Mode

Command: IP PIM CRP Cisco prefix

Configures the c-rp-adv message to a format that a Cisco BSR accepts

Command: no IP PIM CRP Cisco prefix

Configures the c-rp-adv message to the standard RFC format

Default: the c-rp-adv message is in standard RFC format. A Cisco BSR does not receive c-rp-adv messages whose prefix CNT field is 0. By default, if there is no group address, the prefix CNT field is 1 and a 224.0.0.0 group address is filled in. The standard RFC uses 0 in the prefix CNT field when there is no group address. Use this command for compatibility when the BSR is a Cisco router and c-rp-adv messages that a Cisco BSR can recognize must be sent.

28.2.22 Configure Candidate BSR

Mode: Global Configuration Mode

Command: IP PIM BSR candidate < if name > [< hash mask len > | < priority >]

Configures the interface, priority, and hash calculation mask length of the candidate BSR

Command: no IP PIM BSR candidate [if name]

Cancels the candidate BSR identity of the interface

Parameter: < if name > is the interface name of the candidate BSR, which is a Layer 3 interface name (for example, vlan2); < Hash mash len > is the mask length used in the hash algorithm, an integer between 0-32; < Priority > is the priority of the candidate BSR in the BSR election, an integer between 0-255. When the candidate BSR is selected as the BSR, its hash mask len and priority are carried in the bootstrap message.

Default: the default value of priority is 0; the default value of hash mask len is 10.

28.2.23 Configure JP Timer Interval

Mode: Global Configuration Mode

Command: IP PIM JP timer < time >

Configures the upstream join prune timer interval

Command: no IP PIM JP timer [time]

Restores the upstream join prune timer interval to the default time

Parameter: < time > is the interval at which the switch periodically sends the prune message upstream, an integer between 1-65535. The switch maintains the forwarding state of the multicast distribution tree through the JP timer. The holding time of the join and prune state is 3.5 times the JP timer interval, and the holding time is carried in the join prune message.

Default: 60 seconds

28.2.24 Configure SPT Switching

Mode: Global Configuration Mode

Command: IP PIM SPT threshold [group list < ACL Id1 > | < ACL Id2 > | < ACL name >]

Configures SPT switching

Command: no IP PIM SPT threshold [group list < ACL Id1 > | < ACL Id2 > | < ACL name >]

Cancels SPT switching

Parameter: < ACL Id1 > is the number of the standard access control list, an integer between 1-99; < Acl-id2 > is the number of the extended part of the standard access control list, an integer between 1300 and 1999; < ACL name > is the name of the standard access control list. When SPT switching is configured, the access control list can be used to filter the (*, g) entries of the multicast routing table. SPT switching can be configured through the ACL, but it cannot be configured for entries that fail the ACL.

Default: SPT switching is not configured

28.2.25 Configure SSM

Mode: Global Configuration Mode

Command: IP PIM SSM {default | range {< ACL ID > | < ACL name >}}

Configures SSM (source specific multicast)

Command: no IP PIM SSM

Cancels SSM

Parameter: < ACL ID > is the number of the standard access control list, an integer between 1 and 99; < ACL name > is the name of the standard access control list.

Default: SSM is not configured

28.2.26 Configure Multicast Security

Mode: Global Configuration Mode

Command: IP mroute < SRC addr > < GRP addr > < IIF name > < OIF name > remove < if name >

For a multicast route (s, g), stops a specific Layer 2 port in the VLAN that corresponds to an output interface from forwarding the data flow

Command: IP mroute < SRC addr > < GRP addr > < IIF name > < OIF name > Add < if name >

Adds an output port of a multicast route (s, g) back to the VLAN that corresponds to its output interface

Parameters: SRC addr is the multicast source address of the multicast route (s, g)

GRP addr is the multicast group address of the multicast route (s, g)

IIF name is the input interface of the multicast route (s, g)

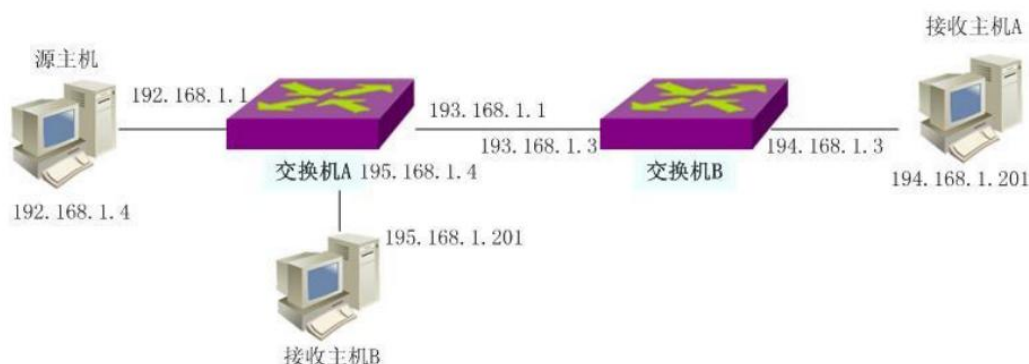
OIF name is the output interface of the multicast route (s, g)

If name is the Layer 2 port, in the VLAN that corresponds to the output interface of the multicast route (s, g). The remove command shields the Layer 2 port from the VLAN where the output interface is located; that is, the data flow forwarded from the output interface is not forwarded from that Layer 2 port. The add command adds the Layer 2 port back to the VLAN where the output interface is located; that is, the data flow forwarded from the output interface is forwarded from that Layer 2 port.

Default: all Layer 2 ports in the VLAN that corresponds to the output interface forward multicast data streams

28.3 PIM-SM Configuration Example

(1) Configuration



Start the PIM-SM protocol so that the receiving host can receive the multicast data stream from the source host.

On switch A:

```
Switch#configure terminal
Switch(config)#ip multicast-routing
Switch(config)#ip pim rp-address 193.168.1.3
Switch(config)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.1/24
Switch(config-vlan2)#ip pim sparse-mode
Switch(config-vlan2)#interface vlan3
Switch(config-vlan3)#ip address 193.168.1.1/24
Switch(config-vlan3)#ip pim sparse-mode
Switch(config-vlan3)#interface vlan5
Switch(config-vlan5)#ip address 195.168.1.1/24
Switch(config-vlan5)#ip pim sparse-mode
```

On switch B:

```
Switch#configure terminal
Switch(config)#ip multicast-routing
Switch(config)#ip pim rp-address 193.168.1.3
Switch(config)#interface vlan3
Switch(config-vlan3)#ip address 193.168.1.3/24
Switch(config-vlan3)#ip pim sparse-mode
Switch(config-vlan3)#interface vlan4
Switch(config-vlan4)#ip address 194.168.1.3/24
Switch(config-vlan4)#ip pim sparse-mode
```

(2) Verification

Use the following command to view PIM-SM information:

```
show ip pim sparse-mode interface
show ip pim sparse-mode neighbor
show ip pim sparse-mode local-members
show ip pim sparse-mode mroute
show ip pim sparse-mode rp mapping
```

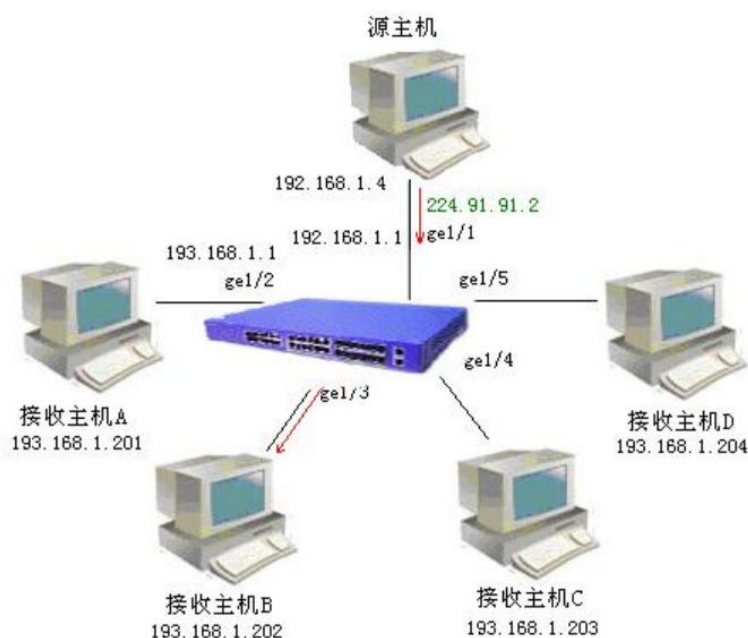
(3) Configure multicast security

The multicast security function module runs on top of the multicast routing protocol. First start the multicast routing protocol, then configure the forwarding state of the Layer 2 ports of its output interface according to the specific origin of a multicast path. In the following figure, vlan2 connects to the multicast source 192.168.1.4. There are four hosts in vlan3, connected to ports GE1 / 2, GE1 / 3, GE1 / 4 and GE1 / 5 respectively. As the red arrow in the figure shows, only the host 193.168.202 requests the programs of multicast group 224.91.91.2 on demand. The other three hosts do not participate in multicast on demand, so the data flow forwarded from those three Layer 2 ports is unnecessary. GE1 / 2, GE1 / 4 and GE1 / 5 are shielded by configuring static multicast commands.

On the switch:

```
Switch#configure terminal
Switch(config)#ip multicast-routing
Switch(config)#ip pim rp-address 193.168.1.1
Switch(config)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.1/24
Switch(config-vlan2)#ip pim sparse-mode
Switch(config-vlan2)#interface ge1/1
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#interface vlan3
Switch(config-vlan3)#ip address 193.168.1.1/24
Switch(config-vlan3)#ip pim sparse-mode
Switch(config-vlan3)#interface ge1/2
Switch(config-ge1/2)#switchport access vlan 3
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#switchport access vlan 3
Switch(config-ge1/3)#interface ge1/4
Switch(config-ge1/4)#switchport access vlan 3
Switch(config-ge1/4)#interface ge1/5
Switch(config-ge1/5)#switchport access vlan 3
Switch(config-ge1/5)#exit
Switch(config)#ip mroute 192.168.1.4 224.91.91.2 vlan2 vlan3 remove ge1/2
Switch(config)#ip mroute 192.168.1.4 224.91.91.2 vlan2 vlan3 remove ge1/4
Switch(config)#ip mroute 192.168.1.4 224.91.91.2 vlan2 vlan3 remove ge1/5
```


Chapter 29 Configuring RIP



This chapter includes the following contents:

- RIP Introduction
- RIP Configuration
- Sample RIP Configuration

29.1 RIP Introduction

RIP (Routing Information Protocol) is an early dynamic routing protocol. It uses a distance vector algorithm and is mostly used in small networks. RIP protocol packets are encapsulated in UDP packets and use UDP port 520. The main idea of RIP is to measure the distance to a host by hop count. The hop count increases by 1 each time the packet passes through a router, which gives the weight (metric) of the route. The RIP convention sets a maximum hop count of 15; a hop count of 16 is marked as unreachable. RIP broadcasts the entire routing table to synchronize the routing information of the routers in the network, and it sends an update message every 30 seconds. If a route entry receives no update message from a neighbor within 180 seconds, it is marked as unreachable. If no valid update is received within 120 seconds, the route is deleted.

RIP is easy to implement because the idea is simple, but it also brings a routing loop problem. To prevent routing loops, RIP introduces the split horizon mechanism, which avoids spoofing between routers. Split horizon means that routing updates are not published out the receiving interface. Split horizon with poison reverse publishes routing updates out the receiving interface, but with the metric marked unreachable, which lets neighboring routers identify loops quickly instead of waiting for the metric to increase to unreachable.

A route entry in the routing table contains a destination address (host or network), a next-hop address, a forwarding interface, a routing metric, a timer that is reset when a routing update is received, and a routing tag.

When RIP starts, it immediately sends a full table request message as a broadcast (RIP-1) or a multicast (RIP-2). The adjacent router receives the request message and sends back its own complete routing table as a response message. When the router receives the response message, it processes the routes one by one and modifies its own routing table. When there is a new route, it immediately generates a trigger update message. After a series of updates, RIP converges, and each router in the network maintains the latest, consistent routing information. After the network is stable, RIP still broadcasts the local routing table to its neighbors every 30 seconds, and each router maintains its own routing information from the received routing update messages to select the most preferred route. RIP uses a time-out mechanism to handle routing entries that have not been updated for a long time, which keeps the route correct in real time.

RIP is mostly used in campus networks and in relatively continuous regional networks with a simple structure. It is not suitable for large, complex networks.

29.2 RIP Configuration

After you start the RIP protocol, you can configure each RIP function and attribute. RIP is mostly configured in RIP configuration mode and interface configuration mode.

The configuration of RIP includes:

- Start RIP and enter RIP configuration mode
- Enable the RIP interface
- Configure unicast message transmission
- Configure the working status of the interface
- Configure the default route weight
- Configure the administrative distance
- Configure the timers
- Configure the version
- Introduce external routes
- Configure route filtering
- Configure an additional routing metric
- Configure the RIP version of the interface
- Configure the receiving and sending status of the interface
- Configure split horizon
- Message authentication
- Configure the interface weight

29.2.1 Start RIP and Enter RIP Configuration Mode

Mode: Global Configuration Mode

Command: router rip

Starts RIP and enters RIP configuration mode

Command: no router rip

Turns off the RIP protocol

Default: the RIP protocol does not run

29.2.2 Enable RIP Interface

When RIP runs, you can designate certain interfaces and configure the network where those interfaces are located as a RIP network, so that RIP protocol messages can be sent and received on the RIP network.

Mode: RIP Configuration Mode

Command: network <network-address>

Enables the RIP interface

Command: no network <network-address>

Shuts down the RIP interface

Parameters: there are two forms, A.B.C.D/M and A.B.C.D A.B.C.D. The former specifies the network IP and mask length; the latter specifies the network IP and mask.

Default: RIP is disabled on all interfaces when it starts

When you start the RIP protocol, you must specify the network segment on which it operates. RIP can run only on interfaces that are on the specified network segment. On interfaces that are not on the specified network segment, RIP neither receives nor forwards routes. Interfaces that are not in the specified network segment do not exist in the RIP view. The parameter network-address is the network address to enable or disable, and it can be configured as the interface IP address. The network command enables the interface of the network segment of that address. For example, if the IP address of an interface is 192.160.1.1, use the command network 192.160.1.1/24; the show running-config command then shows network 192.160.1.0/24.

29.2.3 Configure Unicast Message Transmission

RIP version 1 uses broadcasts to exchange messages, and version 2 uses multicasts (224.0.0.9) to exchange messages. When RIP runs on a link that does not support broadcasts, you must specify a specific unicast address to exchange messages.

Mode: RIP Configuration Mode

Command: neighbor <ip-address>

Configures the peer unicast IP address

Command: no neighbor <ip-address>

Cancels the peer unicast IP address setting

Parameters: ip-address is the specified unicast IP address

Default: RIP does not send messages to any unicast address

29.2.4 Configure the Working Status of the Interface

The RIP protocol operates in some networks where you may need only the RIP interface routes and do not want to broadcast RIP routes on that interface. Use the network command to send and receive RIP protocol messages on the interface and learn the route of the interface. Use the passive-interface command to learn the route of that interface only and block broadcasts on it.

Mode: RIP Configuration Mode

Command: passive-interface <if-name>

Configure the interface as passive.

Command: no passive-interface <if-name>

Cancel the passive state of the interface.

Parameter: if-name is the agreed name of the Layer 3 interface (for example: vlan1, vlan2 ...)

Default: none of the enabled RIP interfaces are passive

29.2.5 Configure the Default Route Weight

When an external route is introduced, a route metric is specified. When it is not specified, the default route metric is used.

Mode: RIP Configuration Mode

Command: default-metric <metric>

Sets the default route metric used when importing an external route

Command: no default-metric [metric]

Restores the default route metric for external routes to the default value of 1

Parameter: the metric value is between 1 and 16, greater than 1 and less than 16.

Default: the metric value is 1. Use the no default-metric command to return to the default value.

29.2.6 Configuration Management Distance

Each protocol has an agreed priority, and the administrative distance is the priority with which a route is selected when a routing policy is used. When there are two identical routes to the same destination, from different routing protocols, the route of the protocol with the smaller administrative distance is preferred.

Mode: RIP Configuration Mode

Command: distance <distance>

Sets the administrative distance value

Command: no distance [distance]

Restore the administrative distance to the default.

Parameter: distance is between 1 and 255

Default: the distance value is 120. Use the no distance command to return to the default value.

29.2.7 Configure the Timer

The RIP protocol has three timers. With the first, the full routing table is broadcast to all RIP interfaces every 30 seconds. With the second, each route in the RIP routing table is marked with a metric of 16 if it receives no update for 180 seconds. With the third, each route in the RIP routing table is deleted from the routing table if it has not been effectively updated for 120 seconds after its metric is 16.

Mode: RIP Configuration Mode

Command: timers basic <update> <timeout> <garbage>

Sets the three timer values

Command: no timers basic

Restores the timers to their defaults

Parameters: the first parameter, update, is the timer for regularly updating the entire RIP routing table; the second parameter, timeout, is the timer after which a route that is not updated times out; the third parameter, garbage, is the timer after which a route is deleted once it has been marked as invalid. The value range of the three timers is 5~(231-1).

Default: update is once every 30 seconds; timeout is 180 seconds, after which the route is marked as invalid; garbage deletes the route in 120 seconds.

29.2.8 Configuration Version

The RIP protocol currently has version 1 (RFC1058) and version 2 (RFC2453). The configured version value is reflected in the version field of the protocol message.

Mode: RIP Configuration Mode

Command: version <version>

Sets the RIP protocol to version 1 or version 2

Command: no version [version]

Restores the RIP protocol version to the default

Parameter: version can take the value 1 or 2

Default: Version 2

29.2.9 Introducing External Routing

RIP lets you import routing information from other protocols into the RIP routing table. The routing protocols (types) that RIP can import include: connected, static, OSPF, IS-IS, and BGP.

Mode: RIP Configuration Mode

Command : redistribute {kernel | connected | static | ospf | isis | bgp} [metric <metric> | route-map <route-map-name>]

Introduces routes of other protocols

Command: no redistribute {kernel | connected | static | ospf | isis | bgp} [metric <metric> | route-map <route-map-name>]

Cancels the imported routes

Parameters: the first parameter is the name of the other protocol to be introduced, including direct connection, static, OSPF, IS-IS, and BGP; the second parameter is the metric value set during introduction, ranging from 1 to 16; the third parameter is the name of the referenced route-map, which is configured in global configuration mode. Refer to the command manual.

Default: the RIP protocol does not introduce any external protocol

29.2.10 Configure Route Filtering

RIP provides a route filtering function. You configure policy rules to filter the received routes and the published routes through a specified access control list and address prefix list.

Mode: RIP Configuration Mode

Command: distribute-list <acl-name> {in | out} [if-name]

Uses an access-list to filter the input and output of the interface

Command: no distribute-list <acl-name> {in | out} [if-name]

Cancels access-list filtering

Parameters: acl-name is the name of the referenced access-list; if-name is the RIP interface to which it is applied; in and out indicate whether it is applied in the direction in which the route is received or in the direction in which the route is published.

Command: distribute-list prefix <pre-name> {in | out} [if-name]

Filters with a prefix-list

Command: no distribute-list prefix <pre-name> {in | out} [if-name]

Cancels prefix-list filtering

Parameters: pre-name is the name of the referenced prefix-list; if-name is the RIP interface to which it is applied; in and out indicate whether it is applied in the direction in which the route is received or in the direction in which the route is published.

Default: RIP does not filter any incoming or outgoing routes

The access-list and prefix-list are configured in global configuration mode; see the command manual.

29.2.11 Configure Additional Routing Metric

The additional routing metric is an offset value added to the routing metric of the RIP protocol on input or output. It does not change the metric of the route in the routing table directly; it adds an offset when the interface receives and sends the route.

Mode: RIP Configuration Mode

Command: offset-list <acl-name> {in | out} <offset> [if-name]

Uses an access-list to add an offset to the weight of the routes the interface inputs and outputs

Command: no offset-list <acl-name> {in | out} <offset> [if-name]

Cancels the offset added to the weight of the input and output routes

Parameters: acl-name is the name of the referenced access-list; in and out indicate whether it is applied in the input or output direction; offset is the value of the offset, between 0 and 16; if-name is the RIP interface to which it is applied.

Default: the additional weight value of each route is 1 when receiving a message, and 0 when sending a message.

29.2.12 Configure the RIP Version of the Interface

RIP has two versions, RIP-1 and RIP-2, and you can specify the version of RIP message that an interface running the RIP protocol processes. In the receiving direction, the interface can receive only RIP-1 messages, receive only RIP-2 messages, or receive both RIP-1 and RIP-2 messages. In the sending direction, it can send RIP-1 messages, send RIP-2 messages (in broadcast mode), send RIP-2 messages (in multicast mode), or send both RIP-1 and RIP-2 messages. RIP-2 has two ways to send messages: broadcast and multicast. Multicast keeps hosts on the same network that do not run RIP from receiving the broadcast messages of RIP, and it also keeps hosts that run RIP-1 from processing RIP-2 routes with a subnet mask incorrectly.

Mode: Interface configuration mode

Command: ip rip receive version {1 | 2}

Sets the interface to receive only version 1 messages or only version 2 messages

Parameter: Version 1 or Version 2

Command: ip rip receive version {1 2 | 2 1}

Sets the interface to receive both version 1 and version 2 messages

Parameter: can be written as 1 2 or 2 1

Command: no ip rip receive version [1 | 2 | 1 2 | 2 1]

Restores the message receiving setting of the interface to the default value

Default: Version 2 multicast mode

Command: ip rip send version {1 | 2 | 1-compatible}

Sets the interface to send only version 1 messages or only version 2 messages

Parameter: version 1 or version 2; 1-compatible means that a version 2 interface sends a message compatible with version 1, that is, a broadcast message rather than a multicast message.

Command: ip rip send version {1 2 | 2 1}

Sets the interface to send both version 1 and version 2 messages

Parameter: can be written as 1 2 or 2 1

Command: no ip rip send version [1 | 2 | 1-compatible | 1 2 | 2 1]

Restores the message sending setting of the interface to the default value.

Default: Version 2 multicast mode

29.2.13 Configure the Receiving and Sending State of an Interface

After you enable the RIP interface with the network command in RIP mode, you can also specify in interface mode the status of receiving and sending protocol messages: whether to receive protocol messages, and whether to send protocol messages.

Mode: Interface configuration mode

Command: ip rip receive-packet

Configures the interface to receive protocol messages

Command: no ip rip receive-packet

Configures the interface not to receive protocol messages

Command: ip rip send-packet

Configures the interface to send protocol messages

Command: no ip rip send-packet

Configures the interface not to send protocol messages

Default: receiving and sending protocol messages is enabled

Note the difference. The network command enables a network to run the RIP protocol. The interface in the network sends and receives protocol messages, and the interface route is included in the routing table. After the network command takes effect, the passive-interface command prevents the interface from sending and receiving protocol messages, but the interface route is still included in the routing table. The IP rip receive-packet and IP rip send-packet commands also specify whether the interface receives or sends protocol packets after the network command takes effect.

29.2.14 Configure Split Horizon

Split horizon means that a route received on an interface is not sent out that interface. Split horizon with poison reverse means that the route received on an interface is still sent out that interface, but its metric value is marked as 16. Split horizon can avoid loops to some extent. Split horizon with poison reverse is more efficient than normal split horizon, because it marks the route unreachable directly. On NBMA networks, however, you must disable split horizon to get the correct route.

Mode: Interface configuration mode

Command: `ip rip split-horizon [poisoned]`

Enables split horizon on the interface, with or without poison reverse

Command: `no ip rip split-horizon`

Disables the split horizon feature on the interface

Parameters: without the poisoned parameter, the normal split horizon function is enabled; with the poisoned parameter, split horizon with poison reverse is enabled.

Default: Split horizon with poison reverse

29.2.15 Message Authentication

RIP-1 does not support message authentication. RIP-2 supports message authentication, with two authentication methods: plaintext authentication and MD5 authentication. In plaintext authentication the authentication data is transmitted unencrypted with the message, which provides no security guarantee and cannot be used on a network with high security requirements. There are two kinds of password settings: the common key and the key chain. The common key stores an independent string; the key chain manages the ID, content, receiving lifetime, and sending lifetime of the key. Key chain management is described in detail in the command reference manual.

Mode: Interface configuration mode

Command: `ip rip authentication mode {text | md5}`

Sets the authentication mode to plaintext or MD5

Command: `no ip rip authentication mode {text | md5}`

Cancels authentication

Parameter: text is plaintext authentication; md5 is MD5 authentication.

Default: No authentication

Command: `ip rip authentication string <password>`

Sets the password string used for authentication

Command: `no ip rip authentication string [password]`

Cancels the authentication password string

Parameter: a 16-byte authentication password

Command: ip rip authentication key-chain <key-chain-name>

Sets the key-chain used for authentication

Command: no IP rip authentication key-chain key-chain-name] Cancels the authentication key-chain

Parameter: the name of the referenced key-chain. The key-chain is configured in global configuration mode. See the command manual.

29.2.16 Configure Interface Weight

Mode: Interface configuration mode

Command: ip rip metric <metric>

Configure the interface weight

Command: no ip rip metric

Restores the interface weight to the default value

Parameter: the metric value is between 1 and 16. It is the metric added to the routing entries that the interface learns.

Default: 1

29.2.17 Display Information

Mode: Normal or Privileged

Command: show ip protocols

Displays information about all protocols that are running

Command: show ip protocols rip

Displays RIP protocol information

Command: show ip rip

Displays RIP routes

Command: show ip rip database

Displays the RIP database

Command: show ip rip database count

Displays the number of RIP database entries

Command: show ip rip interface [if-name]

Displays RIP interface information

Parameter: if-name is the agreed Layer 3 interface name.

Mode: Privileged Mode

Command: show running-config

Displays the current configuration of the switch, including the RIP configuration.

Command: show running-config rip

Displays the current configuration of the RIP protocol.

29.3 Sample RIP Configuration

(1) Configuration

The three switches are connected in pairs and have six network segments respectively, and the RIP protocol is enabled, so that the three PCs can communicate with each other in pairs.

On Switch 1:

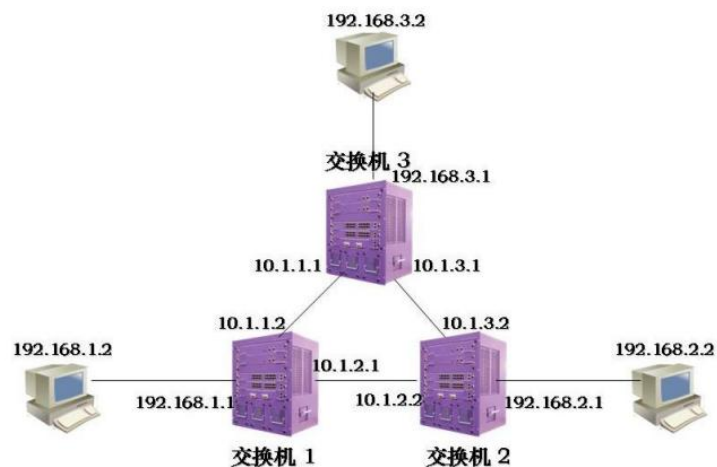
```
Switch# configure terminal
Switch(config)#router rip
Switch(config-rip)#network 192.168.1.0/24
Switch(config-rip)#network 10.1.1.0/24
Switch(config-rip)#network 10.1.2.0/24
```

On Switch 2:

```
Switch# configure terminal
Switch(config)#router rip
Switch(config-rip)#network 192.168.2.0/24
Switch(config-rip)#network 10.1.2.0/24
Switch(config-rip)#network 10.1.3.0/24
```

On Switch 3:

```
Switch# configure terminal
Switch(config)#router rip
Switch(config-rip)#network 192.168.3.0/24
Switch(config-rip)#network 10.1.1.0/24
Switch(config-rip)#network 10.1.3.0/24
```



(2) Verification

Use the following command to view RIP information:

```
show ip protocols rip
show ip rip database
show ip rip interface
```

Chapter 30 Configuring RIPng

This chapter includes the following contents:

- RIPng Introduction
- RIPng Configuration
- RIPng Configuration Example

30.1 RIPng Introduction

RIPng is a relatively simple interior gateway protocol; it is the application of RIP in an IPv6 network. RIPng is mainly used in smaller networks, such as campus networks and regional networks with a simple structure. Because RIPng is relatively simple to implement, and much easier to configure and maintain than OSPFv3 and IS-IS for IPv6, it is still widely used in practical networking.

As IPv6 networks are built, a dynamic routing protocol is also needed to provide accurate, effective routing information for forwarding IPv6 messages. Therefore, while retaining the advantages of RIP, the IETF modified it for IPv6 networks to form RIPng (RIP next generation). RIPng mainly provides the routing function in an IPv6 network. It is an important routing protocol in IPv6 networks.

Differences between RIPng and RIP

To work in an IPv6 network, RIPng modifies the original RIP protocol as follows:

- RIPng uses UDP port 521 (RIP uses port 520) to send and receive routing information.
- The destination address of RIPng uses a 128 bit prefix length (mask length).
- RIPng uses a 128 bit IPv6 address as the next hop address.
- RIPng uses the link local address fe80::/10 as the source address when sending a RIPng routing information update message.
- RIPng sends routing information periodically by multicast, and uses ff02::9 as the router multicast address within the local scope of the link.
- A RIPng message consists of a header and multiple route table entries. In one RIPng message, the maximum number of RTEs is determined by the MTU value of the interface.

30.2 RIPng Configuration

After you start the RIPng protocol, you can configure the various functions and attributes of RIPng. RIPng is mostly configured in RIPng configuration mode and interface configuration mode.

- RIPng configuration includes:
 - Starts RIPng and enters RIPng configuration mode
 - Enables the RIPng interface
 - Configure specific neighbors to send updates
 - Configure the working state of the interface

- Configure the default routing weight
- Configure the timers
- Introduce external routes
- Configure route filtering
- Configure additional routing weights
- Configure split horizon
- Configure interface weights

30.2.1 Start RIPng and Enter RIPng Configuration Mode

Mode: Global Configuration Mode

Command: router IPv6 rip

Starts RIPng and enters RIPng configuration mode

Command: no router IPv6 rip

Closes the RIPng protocol

Default: the RIPng protocol does not run

30.2.2 Enable RIPng Interface

When RIPng runs, you can specify certain interfaces and configure the network they are in as a RIPng network, on which RIPng protocol messages can be sent and received.

Mode: Interface Configuration Mode

Command: IPv6 Router rip

Enables the RIPng interface

Command: no IPv6 Router rip

Closes the RIPng interface

Parameter: none.

Default: RIPng is disabled on all interfaces after the RIPng protocol is started

After you start RIPng, you must specify the interface it works on. RIPng runs only on the specified interface.

30.2.3 Configure Specific Neighbors to Send Updates

Specify neighbors and send updates to those neighbors.

Mode: RIPng configuration mode

Command: neighbor < neighbor address > < ifname >

Configure the IPv6 address of a specific neighbor.

Command: neighbor < neighbor address > < ifname >

Cancel the specific neighbor configuration.

Parameter: neighbor address is the IPv6 address of the specific neighbor.

30.2.4 Configure the Working State of the Interface

On some networks running RIPng, you may need the interface to learn RIPng routes without broadcasting RIPng routes on it. Use the passive interface command so the interface only learns routes and its broadcasts are blocked.

Mode: RIPng configuration mode

Command: passive interface <if name>

Configure the interface as passive.

Command: no passive interface <if name>

Cancels the passive state of the interface.

Parameter: if name is the agreed Layer 3 interface name (for example: vlan1, vlan2...).

30.2.5 Configure Default Routing Weights

When you introduce an external route, you must specify a route weight. If you do not specify one, this default route weight is used.

Mode: RIPng configuration mode

Command: default metric <metric>

Set the default route weight used when introducing external routes.

Command: no default metric [Metric]

Restore the default route weight for external routes to 1.

Parameter: the metric value is between 1 and 16.

Default: the metric value is 1. Use the no default metric command to restore the default value.

30.2.6 Configure Timer

RIPng has three timers. The first broadcasts the complete routing table to all RIPng interfaces every 30 seconds. The second marks a route in the RIPng routing table with a metric of 16 if the route receives no update within 180 seconds. The third deletes a route from the RIPng routing table if the route is marked with a metric of 16 and receives no valid update within 120 seconds.

Mode: RIPng configuration mode

Command: timers basic < update > < timeout > < garbage >

Sets the three timer values

Command: no timers basic

Restore the timers to their default values.

Parameters: the first parameter, update, is the timer for regularly updating the entire RIPng routing table. The second parameter, timeout, is the timer for each route to time out without an update. The third parameter, garbage, is the timer after which each route is deleted once it is marked invalid. The value range of the three timers is 5 ~ (231-1).

Default: update every 30 seconds; a route is marked invalid when timeout reaches 180 seconds; garbage deletes the route in 120 seconds.

30.2.7 Introducing External Routing

RIPng lets you introduce the routing information of other protocols into the rip routing table. The routing protocols (types) that RIPng can introduce are connected, static, OSPF, IS-IS, and BGP.

Mode: RIPng configuration mode

Command: redistribute {kernel | connected | static | ospf6 | Isis | BGP} [Metric < metric > | route map < route map name >]

Introduce routes from other protocols.

Command: no redistribute {kernel | connected | static | ospf6 | Isis | BGP} [Metric < metric > | route map < route map name >] Cancel the introduced route.

Parameter: the first parameter is the name of the other protocol to import: direct connection, static, OSPF, IS-IS, or BGP. The second parameter is the weight set during the import, with a value between 1 and 16. The third parameter is the name of the referenced route map. Route map is configured in global configuration mode; see the command manual.

Default: RIPng does not introduce any external protocols.

30.2.8 Configure Routing Filtering

RIPng provides route filtering. It filters received routes and published routes, and configures policy rules through the specified access control list and address prefix list.

Mode: RIPng configuration mode

Command: distribute list < ACL name > {in | out} [if name]

Use an access list to filter the input and output of the interface.

Command: no distribute list < ACL name > {in | out} [if name]

Cancel access list filtering.

Parameter: ACL name is the name of the referenced access list. If name is the RIPng interface it applies to. in and out indicate whether it applies to the direction of receiving routes or publishing routes.

Command: distribute list prefix < pre name > {in | out} [if name]

Filter using a prefix list.

Command: no distribute list prefix < pre name > {in | out} [if name]

Cancel prefix list filtering.

Parameter: pre name is the name of the referenced prefix list. If name is the RIPng interface it applies to. in and out indicate whether it applies to the direction of receiving routes or publishing routes.

Default: RIPng does not filter any received or sent routes.

Access list and prefix list are configured in global configuration mode. See the command manual.

30.2.9 Configure Additional Routing Weight

The additional routing weight is an offset added to the RIPng route weight on input and output. It does not change the weight of the route in the routing table directly; it adds an offset when the interface receives and sends the route.

Mode: RIPng configuration mode

Command: offset list < ACL name > {in | out} < offset > [if name]

Use an access list to add an offset to the weight of routes the interface receives and sends.

Command: no offset list < ACL name > {in | out} < offset > [if name]

Cancel the offset applied to the weight of received and sent routes.

Parameter: ACL name is the name of the referenced access list. in and out indicate whether it applies to the input or output direction. offset is the value of the offset, which is between 0 and 16. If name is the RIPng interface it applies to.

Default: when a message is received, the additional weight of each route is 1; when a message is sent, the additional weight of each route is 0.

30.2.10 Configure Split Horizon

Split horizon means that a route received on an interface is not sent out that interface. Split horizon with poison reverse means that the route received on the interface is still sent out that interface, but its metric value is marked as 16. Split horizon can avoid loops to a certain extent. Split horizon with poison reverse is more efficient than ordinary split horizon, because it marks the route unreachable directly. However, on an NBMA network you must disable split horizon to obtain the correct route.

Mode: Interface Configuration Mode

Command: IPv6 rip split horizon [poisoned]

Enable split horizon on the interface, or split horizon with poison reverse. Command: no IPv6 rip split horizon

Disable split horizon on the interface.

Parameters: without the poisoned parameter, ordinary split horizon is enabled; with the poisoned parameter, split horizon with poison reverse is enabled.

Default: split horizon with poison reverse

30.2.11 Configure Interface Weights

Mode: Interface Configuration Mode

Command: IPv6 rip metric offset < metric >

Configure the interface weight

Command: no IPv6 rip metric offset < metric >

Restore the interface weight to the default value.

Parameter: the metric value is between 1-16 and indicates the weight added to the routing entries learned on the interface.

Default: 1

30.2.12 Display Information

Mode: normal mode or privileged mode

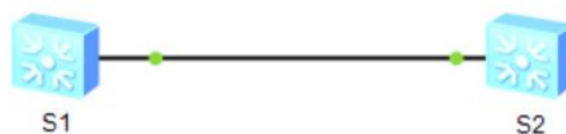
Command: Show IPv6 rip

Display information about RIPng on the current system.

Command: Show IPv6 route rip

Display the currently active RIPng routing information.

30.3 RIPng Configuration Example



Configure switch S1 as follows.

```
Switch(config)#
Switch(config)#int vlan1
Switch(config-vlan1)#ipv6 address 3ffe:506::1/64
Switch(config-vlan1)#exit
Switch(config)#router ipv6 rip
Switch(config-router)#exit
Switch(config)#int vlan1
Switch(config-vlan1)#
Switch(config-vlan1)#ipv6 router rip
Switch(config-vlan1)#
```

Configure switch S2 as follows.

```
Switch(config)#  
Switch(config)#int vlan1  
Switch(config-vlan1)#ipv6 address 3ffe:506::2/64  
Switch(config-vlan1)#exit  
Switch(config)#router ipv6 rip  
Switch(config-router)#exit  
Switch(config)#int vlan1  
Switch(config-vlan1)#  
Switch(config-vlan1)#ipv6 router rip  
Switch(config-vlan1)#
```

Chapter 31 Configuring OSPF

This chapter includes the following contents:

- OSPF introduction
- OSPF configuration
- Sample OSPF configuration

31.1 OSPF Introduction

OSPF (Open Shortest Path First) is a protocol based on the link-state algorithm. It supports large-scale networks and converges quickly.

Each router running OSPF maintains a link-state database (LSDB) that describes the topology of the entire autonomous system, like a map. When the databases of all routers are synchronized, each router calculates the shortest route from its own perspective to every other destination node in the autonomous system and maintains it in its own routing table. When the network topology changes, the router only needs to encapsulate the changed link state in a link state update (LSU) message and broadcast it; all routers then synchronize their local databases again and recalculate the routes. Each router publishes the link-state broadcast (LSA) it sees and collects the LSAs of the others to form the LSDB that describes the topology of the whole network. Convert the LSDB to a weighted directed graph, and you can use the SPF algorithm to calculate the routing table.

On a broadcast network, each router must broadcast its own state information to the other routers, and many pairwise adjacencies are established. This causes a lot of unnecessary message transmission. For this reason, OSPF specifies a designated router (DR) and a backup designated router (BDR). Routers send their link information to the DR, which collects and collates it and sends it to all routers. This effectively reduces the number of adjacencies between routers on the broadcast network.

OSPF supports five types of protocol messages:

1 The HELLO message is broadcast to neighbors periodically to discover and maintain neighbors and to conduct DR election. It contains some interface attribute values. Some parameters in the HELLO message must be consistent before neighbors can be established.

2 DD message (Database Description). During synchronization, a router uses DD messages to describe its own LSDB, including the head of each LSA. An LSA can be uniquely determined from the LSA head, so the peer router can determine whether it has that LSA; if not, it requests the complete LSA.

3 LSR message (Link State Request). After two routers exchange DD messages, they know which of the peer router's LSAs are missing locally. They then send LSR messages to request the complete LSAs. Only the LSA head is required for the request.

4 LSU message (Link State Update) is a collection of multiple LSAs.

5 LSack message (Link State Acknowledgement) acknowledges the received LSU message to ensure reliable transmission of link information. It confirms with the LSA head.

Router-id concept: the unique identification of a router within an autonomous system.

area: If OSPF runs on a large network, the number of routers increases greatly. This makes the LSDB very large, increases the time needed for synchronization and route calculation, and occupies a large amount of storage space and CPU resources. The larger the network, the more frequent the topology changes, so the network is often in change and the router must spend a lot of time transmitting messages and calculating routes. This needlessly occupies network bandwidth.

Therefore, OSPF introduces the concept of the area and divides routers into different areas. The LSDB is synchronized and routes are calculated only within the area, and routing interaction between areas is handled by the border router (ABR). In this way, the number of routers in an area is limited, the LSDB is limited to a smaller capacity, the time to calculate routes is greatly reduced, and convergence is fast when the topology changes.

The concept of the area effectively groups a large-scale network and gives each area a small-scale routing function. Routes between areas interact on the backbone area (the area with area ID 0). Therefore, all non-backbone areas must be connected to the backbone area, that is, the ABR must have at least one interface connected to the backbone area. If the network plan means that a non-backbone area cannot be connected to the backbone area, you must configure a virtual link to establish a logical path, that is, an ABR in the backbone area and an ABR in the non-backbone area establish a point-to-point link through a transmission area. The inter-domain routing information on the backbone area is then also published to the non-backbone area through the virtual link.

31.2 OSPF Configuration

After you start OSPF, enter OSPF configuration mode to set the corresponding attributes and functions. Most OSPF configuration commands are used in OSPF configuration mode and interface configuration mode.

OSPF configuration includes:

- Start OSPF and enter OSPF mode
- Enable the interface
- Specify a host
- Configure the router ID
- Configure adjacent points
- Prevent the interface from sending messages
- Configure the SPF timer
- Configure the administrative distance
- Introduce external routes
- Configure the network type of the interface
- Configure the interval for sending hello messages
- Configure the failure time of the neighbor router
- Configure the retransmission interval
- Configure the interface delay

- Configure the priority of interfaces in DR election
- Configure the cost of sending messages on the interface
- Configure whether to fill in the MTU value in DD messages sent by the interface
- Configure interface message authentication
- Configure an area virtual link
- Configure area route aggregation
- Configure area message authentication
- Configure a stub area
- Configure an NSSA area
- Configure external route aggregation
- Configure the default metric of the external route

31.2.1 Start OSPF and Enter OSPF Mode

OSPF can run as multiple processes, identified by the process number (process-id). When you start OSPF, you must specify which process number to start. Without a parameter, the process number is 0.

Mode: Global Configuration Mode

Command: `router ospf [process-id]`

Start the OSPF process with process number process-id and enter its mode.

Command: `no router ospf [process-id]`

Shut down the OSPF process with process number process-id.

Parameter: the value of process-id is between 1 and (216-1) and indicates the number of the OSPF process to start. If you omit the process-id parameter, OSPF starts with the process number 0.

Default: OSPF is not running

31.2.2 Enable Interface

The merit of OSPF is that it introduces the idea of layering and divides a complete autonomous system into different areas, to establish a conceptual hierarchical network model. Areas are logical, artificial groupings of routers in an autonomous system. When different interfaces of a router belong to different areas, that is, they span areas, the router is called a border router (ABR). Each network segment that starts OSPF can belong to only one specific area, that is, each interface running OSPF on the router must belong to a specified area. An area is identified by an area number (area-id), and the area with the area number 0 is the backbone area. Routing information between different areas is passed through border routers. Unlike RIP, OSPF must be given the area an interface belongs to when it runs on that interface.

Mode: OSPF Configuration Mode

Command: `network <network-address> area <area-id>`

Specify the interface that runs OSPF in the specified area.

Command: no network <network_address> area <area-id>

Turn off OSPF on a specific interface in a specific area.

Parameter: network-address has two forms, A.B.C.D/M and A.B.C.D A.B.C.D. The former specifies the network IP and mask length, the latter the network IP and mask. area-id also has two forms: A.B.C.D, which uses dotted decimal format, and an integer, which takes values between 0 and (232-1).

Default: the interface is not enabled after OSPF is started

31.2.3 Specified Host

Mode: OSPF Configuration Mode

Command: host <ip-address> area <area-id> [cost <cost>]

Configure a host route.

Command: no host <ip-address> area <area-id> [cost <cost>]

Cancel a host route.

Parameters: ip-address uses the A.B.C.D format to represent a specific host within an area, which is represented as a stub on the router's link. area-id is as described for the network command. cost indicates the cost of the specified link and is optional.

Default: if cost is not configured, the default is 0

31.2.4 Configure the Router ID

The router ID is a 32-bit unsigned integer that uniquely identifies a router in an autonomous system. You can configure the router ID manually to ensure that no two routers in the autonomous system have the same ID. If you do not configure it, the router uses the IP address of the loopback interface; if the loopback has no IP address, the highest address among the IP addresses of the current interfaces is selected as the ID. To ensure that OSPF runs stably, divide up and configure the router IDs manually during network planning.

Mode: OSPF Configuration Mode

Command: ospf router-id <router-id>

Configure the router ID.

Command: no ospf router-id

Cancel the router ID.

Command: router-id <router-id>

Command: no router-id [router-id]

Parameters: router-id uses the A.B.C.D format.

Default: when OSPF is started, the router ID is generated automatically according to these rules: first select the router-id configured by the command; if there is no router-id, select the IP address of the loopback; if there is no router-id, select the highest IP address of the current interface; if there is no router-id, select 0.0.

Both sets of commands have the same function.

31.2.5 Configuration of Adjacent Points

OSPF exchanges protocol messages by multicast, using the multicast address 224.0.0.5 or 224.0.0.6. When OSPF runs on a link that does not support broadcasting, such as NBMA, you must configure it to use unicast to exchange protocol messages. In this case, you can specify the IP address of the peer and the corresponding attribute values manually.

Mode: OSPF Configuration Mode

Command : neighbor <ip-address> [priority <prio> | poll-interval <deadtime> | cost <cost>]

Specify the adjacent points at the opposite end and set their properties.

Command : no neighbor <ip-address> [priority <prio> | poll-interval <deadtime> | cost <cost>]

Cancel the adjacent point and attribute settings for the opposite end.

Parameter: ip-address is the IP address of the opposite end, in the format A.B.C.D. prio is the priority of the opposite end, with a value from 0 to 255. deadtime is the timing after which the opposite end is considered to be down; if the timing is terminated, no hello message is sent to the opposite end, with a value from 1 to (216-1). cost is the cost of the link to the opposite end, with a value between 1 and (216-1). Default: priority is 1 (if it is 0, the peer does not participate in DR election); poll-interval is 120 seconds; cost is 10.

31.2.6 Prohibit the Interface From Sending Messages

In a simple network, where an OSPF interface only represents a network segment between two devices and is used only for data transmission, set the interface to the passive state. This blocks hello messages from being broadcast on the link and does not affect the acquisition of the interface route.

Mode: OSPF Configuration Mode

Command: passive-interface <if-name>

Configure the interface as passive.

Command: no passive-interface <if-name>

Cancel the passive state of the interface.

Parameter: if-name is the name of the Layer 3 interface (for example: vlan1 vlan2 ...).

Default: the interfaces enabled after OSPF is started are not passive

If you designate an interface running OSPF as passive, the directly connected route of the interface can still be published, but OSPF messages on the interface are blocked and the interface cannot establish a neighbor relationship. In some networking cases, this effectively saves network resources.

31.2.7 Configure SPF Compute Times

When the link state database (LSDB) of OSPF changes, the shortest path must be recalculated. If the shortest path is calculated immediately for every change, this takes up a lot of resources and affects the efficiency of the router. Configure the two values delay and hold to adjust the interval of the SPF calculation. This suppresses frequent SPF calculations caused by frequent network changes, so that a large amount of system resources is not occupied at one time and the operating efficiency of the router is not affected.

The SPF calculation has a timer that starts the next calculation according to the dead time. When the timer expires and the SPF calculation must start, the holdback time from the last SPF calculation to this time is recalculated. If the configured holdback time has been exceeded, the configured delay time is used to start the timer. If the configured holdback time has not been exceeded, the configured holdback time is used to calculate the required delay time. If that delay time is less than the configured delay time, the configured delay time is used; otherwise, the calculated delay time is used directly to start the SPF calculation.

Mode: OSPF Configuration Mode

Command: `timers spf <delay> <hold>`

Configure the delay and hold values for the SPF calculation interval.

Command: `no timers spf`

Revert to the default.

Parameters: delay indicates the time required to calculate SPF; hold indicates the suppression time required between two SPF calculations.

Default: delay is 5 seconds; hold is 10 seconds

31.2.8 Configuration Management Distance

Multiple routing protocols can run on a router at the same time, and the administrative distance is used to select from the routing information learned by those protocols. When different protocols find the same route, the one with the smaller administrative distance is preferred.

Mode: OSPF Configuration Mode

Command: `distance <distance>`

Configure the management distance.

Command: `no distance <distance>`

Restore the administrative distance to the default.

Command: `distance ospf {intra-area <distance> | inter-area <distance> | external <distance>}`

Configure different types of administrative distance.

Command: `no distance ospf`

Restore the three types of administrative distance to their default values.

Parameter: the value of distance is between 1 and 255. intra-area is the administrative distance of intra-domain routing, inter-area is the administrative distance of inter-domain routing, and external is the management distance of external routing.

Default: the administrative distance of OSPF is 110; the administrative distance of intra-domain route, inter-domain route and external route is 0.

31.2.9 Introducing External Routing

Multiple dynamic routing protocols can run on a router, and different routing protocols can share routing information. OSPF regards the routes learned by other routing protocols as routes outside the autonomous system, introduced by the autonomous system boundary router (ASBR). You can specify attributes such as the weight and the weight type when you introduce an external route.

There are four types of OSPF routes: intra-domain routes and inter-domain routes, both within the autonomous system; and external routes of type-1 and type-2, both describing routes to destinations outside the autonomous system. A type-1 route is a route from another IGP. OSPF considers its credibility relatively high and comparable with the route weight inside the autonomous system. The cost of this kind of external route is therefore the sum of the cost from the router itself to the ASBR and the cost from the ASBR to the destination. A type-2 route is a route from another EGP. OSPF considers its credibility not very high, and its cost is far greater than that of a route inside the autonomous system, so the two are not comparable. The cost of this kind of external route therefore uses only the cost from the ASBR to the destination and ignores the cost from the router itself to the ASBR.

Mode: OSPF Configuration Mode

Command: redistribute {kernel | connected | static | rip | isis | bgp} [metric <metric> | metric-type <type> | route-map <route-map-name> | tag <tag>] Command: no redistribute {kernel | connected | static | rip | isis | bgp} [metric <metric> | metric-type <type> | route-map <route-map-name> | tag <tag>]

Parameter: the first parameter is required and is the type of external route that can be introduced: direct connection, static, RIP, IS-IS or BGP. The second parameter is the weight value set when the external route is introduced, with a value between 0 and (224-1). The third parameter is the two types of external route introduced, type-1 and type-2, where type-1 is the IGP route and type-2 is the EGP route. The third parameter is the name of the referenced route-map, which is configured in global configuration mode; see the command manual. The fourth parameter is tag, with a value between 0 and (232-1), which is the external routing attribute.

Default: no external routing protocol is introduced

31.2.10 Configure the Network Type of the Interface

OSPF is based on the router itself. Each router describes its adjacent network topology and passes it to the other routers. OSPF divides the network type of an interface link into four types according to the protocol type of the link layer. First, the broadcast type (the link layer protocol is Ethernet, FDDI, etc.). Second, the NBMA non-broadcast multiple access type (the link layer protocol is FR, ATM, HDLC, X.25, etc.). Third, the point-to-multipoint type: no link layer protocol is considered point-to-multipoint by default, so the point-to-multipoint type must be forced by configuring it from another network type; the most common practice is to change a

partially connected NBMA network to a point-to-multipoint network. Fourth, the point-to-point type (the link layer protocol is PPP, LAPB, POS).

On broadcast networks that do not have multiaccess capability, you can configure interfaces as the NBMA type. You can configure interfaces as point-to-multipoint when not all routers in the NBMA network are directly reachable from each other.

The NBMA network agreed in OSPF is fully connected, non-broadcast, and multi-point reachable. A point-to-multipoint network is not necessarily fully connected. NBMA requires DR selection; there is no DR in a point-to-multipoint network. An NBMA network sends messages by unicast to the designated neighbor, and a point-to-multipoint network multicasts them.

Mode: Interface configuration mode

Command: `ip ospf network <type>`

Configure the network type for the interface link.

Command: `no ip ospf network`

Restore the network type of the interface link to the default value.

Parameter: type can be broadcast, non-broadcast, point-to-point, or point-to-multipoint [non-broadcast]. The first type is a broadcast network, the second is a non-broadcast network, that is, NBMA, the third is a point-to-point network, and the fourth is a point-to-multipoint network. A point-to-multipoint network is divided into a broadcast network and a non-broadcast network; a non-broadcast neighbor cannot be discovered automatically and must be specified.

Default: broadcast network

31.2.11 Configure the Hello Message Interval

The Hello message is sent to the neighbor router periodically to discover and maintain the neighbor relationship and to elect the DR and BDR. You can configure the interval of Hello messages manually, but the hello timer interval must be the same on all neighbors in the network. The value of the Hello timer is inversely proportional to the router convergence rate and the network load.

Mode: Interface configuration mode

Command: `ip ospf hello-interval <seconds>`

Configure the interval for the hello timer.

Command: `no ip ospf hello-interval`

Restore the hello timer interval to the default.

Parameter: seconds is between 1 and (216-1) and indicates the time interval between two hello messages.

Default: a hello interval of 10 seconds on broadcast and point-to-point networks; a hello interval of 30 seconds on NBMA and point-to-multipoint networks.

31.2.12 Configure the Failure Time of the Neighbor Router

Mode: Interface configuration mode

Command: ip ospf dead-interval <seconds>

Configure the neighbor expiration time.

Command: no ip ospf dead-interval

Restore the neighbor expiration time to the default value.

Parameter: the value of seconds is between 1 and (216-1). The neighbor is considered invalid if no hello message is received from it within seconds. The dead timer of the neighbor is updated every time a hello message is received.

Default: the neighbor failure time is 40 seconds on broadcast networks and point-to-point networks; the neighbor failure time is 120 seconds on NBMA networks and point-to-multipoint networks. When the network type is modified, the hello interval and the dead interval use the default values.

31.2.13 Configure Retransmission Times

OSPF is a reliable link-state protocol. This is reflected in the fact that the LSU messages it exchanges require an LSU-ack response from the opposite end. When the acknowledgement message is received, the link state update is considered to have been received. If no acknowledgement is received within the retransmission interval, the LSA is retransmitted to the neighbor. You can configure the retransmission interval manually. It must be longer than the round-trip time of a message transmitted between two routers. If you set it too small, it causes unnecessary retransmission.

Mode: Interface configuration mode

Command: ip ospf retransmit-interval <seconds>

Configure the retransmission interval for the interface.

Command: no ip ospf retransmit-interval

Restore the retransmission interval to the default.

Parameter: seconds is between 1 and (216-1) and indicates the retransmission interval used when the peer does not receive the LSA.

Default: the retransmission interval is 5 seconds

31.2.14 Configure Interface Delay

Each link state broadcast (LSA) in a link state update (LSU) message has an age time field, and the transmission delay of the sending interface must be added to it before transmission. This parameter mainly accounts for the time the interface needs to send the message. It is especially necessary to configure this parameter on a low-speed network.

Mode: Interface configuration mode

Command: ip ospf transmit-delay <seconds>

Set the transmission delay for the interface.

Command: no ip ospf transmit-delay

Restore the transmission delay of the interface to the default value.

Parameter: the value of seconds is between 1 and (216-1) and indicates that the age field of an LSA sent on this interface must be increased by this delay value.

Default: The transmission delay of the interface is 1s.

31.2.15 Configure the Priority of Interfaces in DR Election

On a broadcast network, to avoid repeated point-to-point transmission of link information, you must elect a designated router (DR) and a BDR to broadcast the link information in the network segment. The priority of an interface indicates its qualification in the election of a DR. If there is a conflict in the election, the interface with the highest priority is considered first. If the priority is 0, the router does not participate in the election; if the priority is greater than 0, the router is a candidate. Each router includes its own priority information and its own DR in the hello message it broadcasts on the broadcast network. The router with the higher priority is finally selected as the DR. If the priorities are equal, the higher router ID is compared.

After the DR fails, the routers in the network must go through a process of re-electing the DR. This takes a long time and can cause routing calculation errors while it runs. The BDR exists to make the transition to the new DR smooth. The BDR is the backup of the DR and is elected at the same time as the DR. It also establishes an adjacency relationship with the other routers in the network, but the node that collects and publishes information in the network is the DR, not the BDR; the BDR only maintains the synchronization of the adjacency. When the DR fails, the BDR immediately becomes the DR and is responsible for collecting the information in the network segment. A new process then starts to elect the BDR, but the election of the BDR does not affect route calculation.

Mode: Interface configuration mode

Command: ip ospf priority <prio>

Configure the priority of the interface in DR election.

Command: no ip ospf priority

Restore the interface priority to the default.

Parameter: the value of prio is 0 ~ 255 and indicates the priority in DR selection. A value of 0 means the interface does not participate in the election.

Default: priority is 1

31.2.16 Configure the Cost of Sending Messages on an Interface

You control traffic by configuring different costs for different links in the network. The cost of an interface represents the cost of sending messages from that interface. If you do not configure it manually, OSPF calculates the interface cost automatically from the interface baud rate.

Mode: Interface configuration mode

Command: ip ospf cost <cost>

Command: no ip ospf cost

Parameter: the cost value is between 1 and (216-1) and indicates the cost of sending a message on this interface.

Default: interface cost is 10

31.2.17 Configure Whether the Interface Fills the MTU Field in DD Messages

Mode: Interface configuration mode

Command: ip ospf mtu-ignore

Set the MTU value in DD messages not to be checked.

Command: no ip ospf mtu-ignore

Cancel the setting that skips the MTU check in DD messages.

Default: check the MTU value in DD messages

31.2.18 Configure Interface Message Authentication

OSPF message authentication on the interface supports plaintext mode and MD5 mode.

Mode: Interface configuration mode

Command: ip ospf authentication <mode>

Configure the authentication mode.

Command: no ip ospf authentication

Cancels authentication

Parameter: no parameter means plaintext authentication; message-digest means MD5 authentication; null means no authentication.

Command: ip ospf authentication-key <password>

Configure the plaintext authentication password string.

Command: no ip ospf authentication-key

Cancel the plaintext authentication password string.

Parameter: password is the password string for plaintext authentication.

Command: ip ospf message-digest-key <key-id> md5 <password>

Configure the MD5 authentication password.

Command: no ip ospf message-digest-key <key-id>

Cancel the MD5 authentication password.

Parameter: the value of key-id is between 1 and 255 and is used to sort in the key chain; password is the password string.

Default: no authentication is configured

31.2.19 Configure Area Virtual Link

OSPF uses the idea of layering to divide the routers in an autonomous system into different groups. These groups are called areas. The areas are not all equal; they have a hierarchical relationship. The 0.0.0.0 area is the most special of them: it is the backbone area. Other, non-backbone areas must exchange inter-domain routes through the backbone area. Therefore, all non-backbone areas must be connected to the backbone area, that is, at least one interface on the ABR must be in area 0. If some areas cannot be given physical access to the backbone area because of the network topology, you must configure virtual links to ensure logical access. Both ends of a virtual link are ABRs, and the link passes through a non-backbone area in the middle, which is called a transit area. When you configure the virtual link, you must specify the ID of the transmission area and the ID of the ABR at the opposite end. The virtual link must be configured on the ABRs at both ends to take effect.

When the virtual link is activated after the route of the transmission area is calculated, it is logically equivalent to a point-to-point connection between the two endpoints. You can therefore configure the interface parameters on the physical interface and start the authentication function.

What is transmitted between the ABRs is a unicast packet, and the router that forwards the unicast packet in the transmission area forwards it as an ordinary IP packet. It can therefore only be understood as a logical link provided in the transmission area, over which the two ABRs can exchange protocol packets.

Mode: OSPF Configuration Mode

Command: area <area-id> virtual-link <router-id>

Configure the transmission area of the virtual link and the ID of the opposite end.

[authentication <mode> |

Configure the authentication mode of the virtual link.

```
authentication-key <password> |
```

Configure the virtual link plaintext authentication password. message-digest-key <key-id> md5 <password> |

Configure the virtual link MD5 authentication password. hello-interval <seconds> |

Configure the hello interval for the virtual link. dead-interval <seconds> |

Configure the virtual link neighbor failure time. retransmit-interval <seconds> |

Configures the retransmission interval of the virtual link transmit-delay <seconds> |

Configure the virtual link interface delay.

Command: no area <area-id> virtual-link <router-id>

[authentication <mode> |

authentication-key <password> |

message-digest-key <key-id> md5 <password> | hello-interval <seconds> | dead-interval <seconds> |
retransmit-interval <seconds> | transmit-delay <seconds>]

Cancel the virtual link settings.

Parameter: area-id is the ID of the transmission area. It can be A.B.C.D in dotted decimal format or an integer, with a value between 0 and (232-1). router-id is the ID of the router at the opposite end of the virtual link, in A.B.C.D format. The authentication and sending interface attributes are optional; see the related command descriptions.

Default: no virtual link is configured

31.2.20 Configure Area Route Aggregation

Mode: OSPF Configuration Mode

Command: area <area-id> range <ip-prefix> [advertise | not-advertise]

Configure the aggregation range.

Command: no area <area-id> range <ip-prefix> [advertise | not-advertise] Cancel the aggregation.

Parameter: area-id is the area ID and specifies that the routes in that area are aggregated. It can be A.B.C.D in dotted decimal format or an integer, with a value between 0 and (232-1). ip-prefix uses the prefix format A.B.C.D/M to indicate the aggregation range. The optional parameters advertise and not-advertise indicate whether to broadcast the aggregation range, that is, ip-prefix. The original network routes are broadcast.

31.2.21 Configure Area Message Authentication

All routers in an area must have the same authentication type. The authentication password string must be consistent on all routers in a network segment. Configuring area authentication only starts the authentication function (plaintext or MD5); the password uses the corresponding value configured on the interface. See the interface message authentication configuration.

Mode: OSPF Configuration Mode

Command: area <area-id> authentication [message-digest]

Configure the area authentication mode.

Command: no area <area-id> authentication

Cancel area authentication.

Parameter: area-id is the area ID and specifies the area to be authenticated. You can use the dotted decimal format A.B.C.D or an integer, with a value between 0 and (232-1). No optional parameter means plaintext authentication; message-digest means MD5 authentication.

Default: area authentication is not enabled

31.2.22 Configure Stub Area

Mode: OSPF Configuration Mode

Command: area <area-id> stub [no-summary]

Configure the router to be in the stub area.

Command: no area <area-id> stub [no-summary]

Remove the router from the stub area.

Command: area <area-id> default-cost <cost>

Configure the default cost for ABR broadcast routes connected to the stub area.

Command: no area <area-id> default-cost

Restore the default cost to the default value.

Parameter: area-id is the area ID and indicates which area has the stub attribute. You can use A.B.C.D in dotted decimal format or an integer, with a value between 0 and (232-1). no-summary means that no inter-domain route is injected into the stub area.

The first set of commands configures the router in the stub area; the second set configures the ABR whose interface is connected to the stub area.

Default: no stub area is configured

31.2.23 Configure the NSSA Area

Mode: OSPF Configuration Mode

Command: area <area-id> nssa [options]

Configure nssa properties.

Command: no area <area-id> nssa [options]

Cancel the nssa area attribute.

Parameters: area-id is the area ID. For details about options, see the command manual.

Default: the nssa area is not configured

31.2.24 Configure External Route Aggregation

Routes introduced from other protocols are broadcast one by one in type-5 LSUs. Use the aggregation command to specify a prefix range: the routes it covers are suppressed and only the aggregated route is broadcast. When the number of external routes is huge, this effectively reduces the size of the LSDB.

Mode: OSPF Configuration Mode

Command: summary-address <ip-prefix> [not-advertise | tag <tag>]

Configure the aggregation scope and properties.

Command: no summary-address <ip-prefix> [not-advertise | tag <tag>]

Cancel external route aggregation.

Parameter: ip-prefix uses the address prefix format A.B.C.D/M to indicate the route range to aggregate. not-advertise means the aggregated route is not broadcast. tag is the tag value that is set, with a value between 0 and (232-1); the default is 0.

Default: external introduced routes are not aggregated

31.2.25 Configure the Default Metric of the External Route

When external routes are introduced, the default metric is used if the redistribute command does not specify a metric value.

Mode: OSPF Configuration Mode

Command: default-metric <metric>

Configure the default metric used when importing an external route.

Command: no default-metric [metric]

Restore the default weight used when an external route is introduced to its default value.

Parameter: the metric value is between 0 and (224-1).

Default: the default weight value is 1

31.2.26 Display Information

Mode: Normal or Privileged

Command: show ip protocols

Command: show ip protocols ospf

Display OSPF protocol information.

Command: show ip ospf [process-id]

Display OSPF process information.

Parameter: instance-id is the process number.

The value is between 0 and (216-1).

Command: show ip ospf border-routers

Display ABR information.

Command: show ip ospf database <type>

Displays LSDB information.

Parameter: type is each type of LSA and the summary information. See the command manual for details.

Command: show ip ospf interface [if-name]

Display OSPF interface information.

Parameter: if-name is the agreed Layer 3 interface name.

Command: show ip ospf route[count]

Display the OSPF routing table.

Parameter: count displays the total number of entries in the routing table.

Command: show ip ospf virtual-links

Display OSPF virtual connection information.

Command: show ip ospf neighbor [options]

Display OSPF neighbor information.

Parameters: for options, see the command manual for details.

Mode: Privileged Mode

Command: show running-config

Display the current configuration of the switch, including the OSPF configuration.

Command: show running-config ospf

Display the current OSPF configuration.

31.2 Sample OSPF Configuration

(1) Configuration

The three switches are connected in pairs and have six network segments. OSPF is enabled so that the three PCs can communicate with each other. All interfaces must be in the same area, area 0.

On Switch 1:

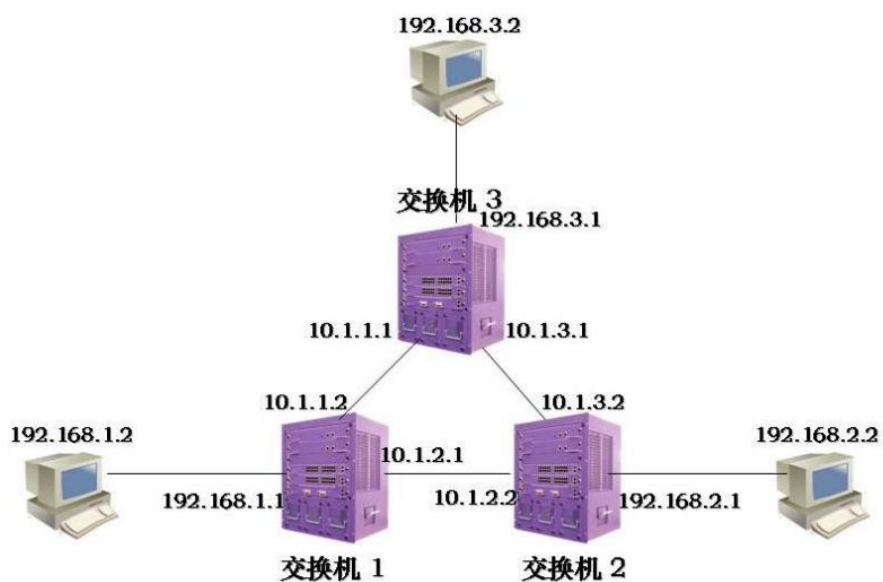
```
Switch#configure terminal
Switch(config)#router ospf 100
Switch(config-ospf-100)#network 10.1.1.0/24 area 0
Switch(config-ospf-100)#network 10.1.2.0/24 area 0
Switch(config-ospf-100)#network 192.168.1.0/24 area 0
```

On Switch 2:

```
Switch#configure terminal
Switch(config)#router ospf 100
Switch(config-ospf-100)#network 10.1.2.0/24 area 0
Switch(config-ospf-100)#network 10.1.3.0/24 area 0
Switch(config-ospf-100)#network 192.168.2.0/24 area 0
```

On Switch 3:

```
Switch#configure terminal
Switch(config)#router ospf 100
Switch(config-ospf-100)#network 10.1.1.0/24 area 0
Switch(config-ospf-100)#network 10.1.3.0/24 area 0
Switch(config-ospf-100)#network 192.168.3.0/24 area 0
```



(2) Verification

```
show ip ospf database
show ip ospf interface
show ip ospf neighbor
show ip route ospf
show ip ospf route
```

Chapter 32 Configuring OSPFv3

This chapter includes the following contents:

- OSPFv3 introduction
- OSPFv3 configuration
- OSPFv3 configuration example

32.1 OSPFv3 Introduction

OSPF (open shortest path first) is an internal gateway protocol based on link state, developed by the IETF Interior gateway protocol.

At present, OSPF version 2 is used for the IPv4 protocol and OSPF version 3 is used for the IPv6 protocol.

- OSPFv3 is short for OSPF version 3.
- OSPFv3 is an OSPF routing protocol running on IPv6 (rfc5340, the same as rfc2740).
- OSPFv3 is modified on the basis of OSPFv2 and is an independent routing protocol.

The main purpose of OSPFv3 is to develop a routing protocol independent of any specific network layer. To achieve this,

the internal router information of OSPFv3 has been redesigned.

OSPFv3 differs from OSPFv2 in that:

- OSPFv3 does not insert IP-based data into the message header at the beginning of the packet or into link state advertisement (LSA) data.
- OSPFv3 uses information independent of the network protocol to perform key services that previously required IP message header data, such as identifying the LSA that publishes routing data.

Basic principle of OSPFv3

OSPFv3 is an OSPF routing protocol (rfc2740) running on IPv6. It is enhanced on the basis of OSPFv2 and is an independent routing protocol.

- The working principle of OSPFv3 is consistent with OSPFv2 in the Hello message, state machine, LSDB, flooding mechanism, route calculation, and so on.
- OSPFv3 divides the autonomous system into one or more logical areas and publishes routes in the form of LSAs (link state advertisements).
- OSPFv3 relies on the exchange of OSPFv3 messages between devices in an OSPFv3 area to unify the routing information.
- An OSPFv3 message is encapsulated in an IPv6 message and can be sent by unicast or multicast.

OSPFv3 message type

Message Type	Message Function
Hello Message	Sent periodically to discover and maintain OSPFv3 neighbor relationships.
DD Message (Database Description packet)	Describes the summary information of the local LSDB, used for database synchronization between two devices.
LSR Message (Link State Request packet)	Used to request the required LSA from the other party. A device sends an LSR message to its OSPFv3 neighbors only after they successfully exchange DD messages.
LSU Message (Link State Update packet)	Sends the required LSA to the other party.
LSAck Message (Link State Acknowledgment packet)	Used to confirm the received LSA.

LSA Type

LSA Type	LSA Function
Router-LSA (Type1)	The device generates an LSA for the area where each OSPFv3 interface runs. It describes the link state and overhead of the device and propagates in the area it belongs to.
Network-LSA (Type2)	Generated by the DR. Describes the link state of the link and propagates in the area it belongs to.
Inter-Area-Prefix-LSA (Type3)	Generated by the ABR. Describes the route of a network segment in the area and notifies other relevant areas.
Inter-Area-Router-LSA (Type4)	Generated by the ABR. Describes the route to the ASBR and notifies other relevant areas, except the area where the ASBR is located.
AS-external-LSA (Type5)	Generated by the ASBR. Describes the route outside the AS and notifies all areas (except the stub area and the NSSA area).
NSSA LSA (Type7)	Generated by the ASBR. Describes the route outside the AS and propagates only in the NSSA area.
Link-LSA (Type8)	Each device generates a link LSA for each link. It describes the link local address and IPv6 prefix address on that link and provides the link options that are set in the network LSA. It propagates only within that link.
Intra-Area-Prefix-LSA (Type9)	Each device and DR generates one or more of these LSAs and propagates them in its area. Such an LSA generated by a device describes the IPv6 prefix address associated with the router LSA; such an LSA generated by the DR describes the IPv6 prefix address associated with the network LSA.

32.2 OSPFv3 Configuration

32.2.1 Start OSPFv3 and Enter OSPFv3 Configuration Mode

OSPFv3 can run as multiple processes, identified by process ID. When you start OSPFv3, you must specify which process number to start. Without a parameter, the process number is 0.

Mode: Global Configuration Mode

Command: router IPv6 OSPF [process ID]

Start the OSPFv3 process with this process ID and enter its mode.

Command: no IPv6 Router OSPF [process ID]

Close the OSPFv3 process with this process ID.

Parameter: the value of process ID is between 1 and 216-1 and indicates the number of the OSPF process to start. If you omit the process ID parameter, OSPFv3 starts with the process number 0.

Default: The OSPFv3 protocol is not running.

32.2.2 Configure Router ID

The router ID is a 32-bit unsigned integer that uniquely identifies a router within the autonomous system. You can configure the router ID manually. When you do, make sure that no two routers in the autonomous system have the same ID. If it is not configured, the router uses the IP address of the loopback interface; if the loopback interface has no IP address, the router selects the highest of the IP addresses of the current interfaces as the ID. To keep OSPFv3 running stably, allocate and configure the router IDs manually during network planning.

Mode: OSPFv3 configuration mode

Command: router ID <router ID>

Configures the router ID.

Command: no router ID

Cancels the router ID.

Command: router ID <router ID>

Command: no router ID [router ID]

Parameter: The router ID uses a.b.c.d format.

Default: After the OSPFv3 protocol starts, the router ID is generated automatically by the following rules: first, use the router ID configured by the command; if there is none, use the IP address of the loopback interface; if there is none, use the highest IP address of the current interface; if there is still none, use 0.0.0.0.

The two groups of commands have the same function.

32.2.3 Prohibit the Interface from Sending Messages

In a simple network, where an OSPFv3 interface represents only a network segment between two devices used for data transmission, set the interface to the passive state. This blocks the Hello message broadcast on its link and does not affect learning of the interface route.

Mode: OSPFv3 Configuration Mode

Command: `passive interface <if name>`

Configure the interface as passive.

Command: `no passive interface <if name>`

Cancels the passive state of the interface.

Parameter: `if name` is the name of the Layer 3 interface (for example, `vlan1`, `vlan2...`).

Default: After the OSPFv3 protocol starts, no enabled interface is in the passive state.

When you specify an interface running OSPFv3 as passive, the direct route of the interface can still be published, but OSPFv3 messages on the interface are blocked and the interface cannot establish a neighbor relationship. In some networking cases this saves network resources effectively.

32.2.4 Configure SPF Calculation Time

When the OSPFv3 link state database (LSDB) changes, the shortest path must be recalculated. Calculating the shortest path immediately on every change occupies many resources and affects the efficiency of the router. By configuring the delay and hold values, you adjust the interval between SPF calculations. This suppresses the overly frequent SPF calculations caused by frequent network changes, so the router does not occupy a large amount of system resources at one time and its operating efficiency is not affected.

SPF calculation uses a timer that starts the next calculation according to the suppression time. When an SPF calculation must start as the timer terminates, the suppression time from the last SPF calculation to this one is recalculated. If the configured suppression time has been exceeded, the configured delay time is used to start the timer. If the configured suppression time has not been exceeded, the configured suppression time is used to calculate the required delay time. If that delay time is less than the configured delay time, the configured delay time is used; otherwise, the calculated delay time is used directly to start the SPF calculation.

Mode: OSPFv3 Configuration Mode

Command: `timers SPF <delay> <hold>`

Configures the delay and hold values of the SPF calculation interval.

Command: `no timers SPF`

Restores the default.

Parameter: `delay` is the delay time required to calculate SPF; `hold` is the suppression time between two SPF calculations.

Default: `delay` is 5S; `hold` is 10 seconds.

32.2.5 Configure Management Distance

A router can run several routing protocols at the same time. The management distance decides how to choose among the routing information learned by the different protocols. When different protocols find the same route, the one with the smaller management distance is preferred.

Mode: OSPFv3 Configuration Mode

Command: distance <distance>

Configures the management distance.

Command: no distance <distance>

Restores the management distance to the default value.

Command: distance OSPFv3 {intra area <distance> | inter area <distance> | external <distance>}

Configures the different types of management distance.

Command: no distance OSPFv3

Restores the three types of management distance to the default value.

Parameter: distance is between 1 and 255; intra area is the management distance of routes within the domain; inter area is the management distance of inter-domain routes; external is the management distance of external routes.

Default: The management distance of the OSPFv3 protocol is 110. The management distance of intra-domain routes, inter-domain routes, and external routes is 0.

32.2.6 Introducing External Routing

A router can run several dynamic routing protocols, and routing information can be shared between them. OSPFv3 treats a route learned by another routing protocol as a route outside the autonomous system, introduced by the autonomous system boundary router (ASBR). When you introduce an external route, you can specify attributes such as the weight and the weight type.

OSPFv3 has four types of routes. The first is intra-domain routing and the second is inter-domain routing; these two types are inside the autonomous system. The third is the type-1 external route and the fourth is the type-2 external route; these two types describe routes to destinations outside the autonomous system. Type-1 routes come from other IGPs. OSPFv3 considers them highly reliable and comparable with route weights inside the autonomous system, so the cost of this kind of external route is the sum of the cost from the router itself to the ASBR and the cost from the ASBR to the destination. Type-2 routes come from other EGP. OSPFv3 considers their reliability lower and their cost far greater than costs inside the autonomous system, so they are not comparable. The cost of this kind of external route therefore uses only the cost from the ASBR to the destination and ignores the cost from the router itself to the ASBR.

Mode: OSPFv3 Configuration Mode

Command: redistribute {kernel | connected | static | rip | Isis | BGP} [Metric <metric> | metric type <type> | route map <route map name> | tag <tag>] Command: no redistribute {kernel | connected | static | rip | Isis | BGP} [Metric <metric> | metric type <type> | route map <route map name> | tag <tag>]

Parameter: The first parameter is required and gives the type of external route that can be introduced: direct connection, static, rip, IS-IS, and BGP. The second parameter is the weight value set when the external route is introduced, which is between 0 and 224-1. The third parameter is the type of external route introduced, either type-1 or type-2; type-1 is an IGP route and type-2 is an EGP route. The third parameter is the name of the referenced route map. The route map is configured in global configuration mode; see the command manual. The fourth parameter is tag, which is between 0 and 232-1. It is an external routing attribute.

32.2.7 Configure the Network Type of the Interface

The OSPFv3 protocol takes the router itself as its perspective. Each router describes its adjacent network topology and passes it to the other routers. OSPFv3 divides the network types of interface links into four types according to the link layer protocol type. The first is the broadcast type (link layer protocols such as Ethernet and FDDI). The second is the NBMA non-broadcast multiple access type (link layer protocols such as FR, ATM, HDLC, and X.25). The third is the point-to-multipoint type; no link layer protocol is treated as point-to-multipoint by default, so this type must be forced by configuring it from another network type. The most common approach is to change a non-fully-connected NBMA network to a point-to-multipoint network. The fourth is the point-to-point type (link layer protocols such as PPP, LAPB, and POS).

On a broadcast network without multiple access capability, you can configure the interface as the NBMA type. When not all routers in an NBMA network are directly reachable, you can configure the interface as the point-to-multipoint type.

The NBMA network agreed in the OSPFv3 protocol is fully connected, non-broadcast, and reachable at multiple points. A point-to-multipoint network is not necessarily fully connected. NBMA requires a DR to be elected; a point-to-multipoint network has no DR. On an NBMA network, multicast messages are sent as unicast messages through designated neighbors; on a point-to-multipoint network, multicast messages are used.

Mode: Interface Configuration Mode

Command: IPv6 ospf network <type>

Configures the network type of the interface link.

Command: no IPv6 ospf network

Restores the network type of the interface link to the default value.

Parameter: type can be broadcast, non broadcast, point to point, or point to multipoint [non broadcast]. The first type is a broadcast network; the second is a non-broadcast network, that is, NBMA; the third is a point-to-point network; the fourth is a point-to-multipoint network. Point-to-multipoint networks are divided into broadcast and non-broadcast networks. Non-broadcast neighbors cannot be found automatically and must be specified.

Default: broadcast network

32.2.8 Configure Hello Message Sending Interval

Hello messages are sent to neighbor routers periodically to find and maintain neighbor relationships and to elect the DR and BDR. You can configure the Hello message interval manually, but take care to keep the Hello timer interval consistent between neighbors in the network. The value of the Hello timer is inversely proportional to the convergence speed of the router and to the network load.

Mode: Interface Configuration Mode

Command: IPv6 OSPF Hello interval <seconds>

Configures the Hello timer interval.

Command: no IPv6 OSPF Hello interval

Restores the Hello timer interval to the default value.

Parameter: seconds is between 1 and 216-1. It is the time interval between the sending of two Hello messages.

Default: On broadcast networks and point-to-point networks, the Hello interval is 10 seconds. On NBMA networks and point-to-multipoint networks, the Hello interval is 30 seconds.

32.2.9 Configure Neighbor Router Expiration Time

Mode: Interface Configuration Mode

Command: IPv6 OSPF dead interval <seconds>

Configures the neighbor expiration time.

Command: no IPv6 OSPF dead interval

Restores the neighbor failure time to the default value.

Parameter: seconds is between 1 and 216-1. If no Hello message is received after this many seconds, the neighbor is considered invalid. Each time a Hello message is received, the neighbor's dead timer is updated.

Default: The failure time of neighbors on broadcast networks and point-to-point networks is 40 seconds. On NBMA networks and point-to-multipoint networks, the neighbor failure time is 120 seconds. When the network type is modified, the Hello interval and dead interval use the default values.

32.2.10 Configure Retransmission Time

OSPFv3 is a reliable link state protocol: the LSU messages it exchanges require an LSU ack from the far end. When the acknowledgement message is received, the sender considers the link state update received. If no acknowledgement message is received within the retransmission interval, the LSA is retransmitted to the neighbor. You can configure the retransmission interval manually. It must be longer than the time a message takes to travel back and forth between two routers. Setting it too small causes unnecessary retransmission.

Mode: Interface Configuration Mode

Command: IPv6 OSPF retransmit interval <seconds>

Configures the retransmission interval of the interface.

Command: no IPv6 OSPF retransmit interval

Restores the retransmission interval to the default value.

Parameter: seconds is between 1 and 216-1. It is the retransmission interval used when the far end does not receive the LSA.

Default: retransmission interval 5S

32.2.11 Configure Interface Delay

In the link state update message LSU, each link state broadcast LSA has a time domain. Before transmission, the transmission delay of the sending interface must be added. This parameter mainly accounts for the time the interface needs to send messages, especially on low-speed networks.

Mode: Interface Configuration Mode

Command: IPv6 OSPF transmit delay <seconds>

Sets the transmission delay of the interface.

Command: no IPv6 OSPF transmit delay

Restores the transmission delay of the interface to the default value.

Parameter: seconds is between 1 and 216-1. It is the delay value that must be added to the age domain of the LSA sent by the interface.

Default: The transmission delay of the interface is 1s.

32.2.12 Configure the Priority of the Interface in DR Election

To avoid repeated point-to-point transmission of link information in a broadcast network, a designated router (DR) and a BDR are elected to broadcast the link information of the network segment. The priority of an interface indicates its qualification in the DR election. In case of an election conflict, the one with the higher priority is considered first. An interface with a priority of 0 does not take part in the election; an interface with a priority greater than 0 is a candidate. Each router includes its own priority information and its own DR in its Hello messages and broadcasts them on the broadcast network, and the one with the higher priority finally becomes the DR. If the priorities are equal, the higher router ID prevails.

When the DR fails, the routers in the network must go through a process of re-electing the DR. This takes time, and during this time routing calculation errors occur. The concept of the BDR is to make a smooth transition to the new DR. The BDR is the backup of the DR and is elected at the same time as the DR. It also establishes an adjacency relationship with the other routers in the network. Only the DR, not the BDR, is the information collection and publishing node in the network. The BDR only maintains the synchronization of the adjacency. When the DR fails, the BDR immediately becomes the DR and is responsible for collecting information in the network segment. A new process then starts to elect the BDR, but the election of the BDR does not affect the calculation of routes.

Mode: Interface Configuration Mode

Command: IPv6 OSPF priority <prio>

Configures the priority of the interface in the DR election.

Command: no IPv6 OSPF priority

Restores the interface priority to the default

Parameter: prio is 0 to 255. It is the priority in DR selection. A value of 0 means the interface does not take part in the election.

Default: priority is 1

32.2.13 Configure the Cost of Sending Messages on an Interface

In the network, traffic is controlled by configuring different costs on different links. The cost of an interface is the cost of sending messages from that interface. If it is not configured manually, OSPFv3 calculates the interface cost automatically from the interface baud rate. Mode: interface configuration mode

Command: IPv6 OSPF cost <cost>

Command: no IPv6 OSPF cost

Parameter: cost is between 1 and 216-1. It is the proxy value of the messages sent on the interface.

Default: interface cost 10

32.2.14 Configure Whether the Interface Fills the MTU Field in DD Messages

Mode: Interface Configuration Mode

Command: IPv6 OSPF MTU ignore

Sets the MTU value in DD messages not to be checked.

Command: no IPv6 OSPF MTU ignore

Cancels the setting not to check the MTU value in DD messages.

Default: The MTU value in DD messages is checked.

32.2.15 Configure Regional Virtual Link

The OSPFv3 protocol uses a layered approach to divide the routers in an autonomous system into different groups. These groups are called regions. The regions are not equal and parallel; they have a hierarchical relationship. Region 0.0.0.0 is the most special: it is the backbone region. Other, non-backbone regions must exchange inter-domain routes through the backbone region. Therefore, all non-backbone areas must be connected to the backbone area, that is, at least one interface on the ABR is in area 0. If some areas cannot be guaranteed a physical path to the backbone area because of network topology limits, you must configure virtual links to guarantee a logical path. Both ends of a virtual link are ABRs, and the link passes through a non-backbone area in the middle, called the transmission area. When you configure a virtual link, you must specify the ID of the transmission area and the ID of the peer ABR, and you must configure it on the ABRs at both ends before it takes effect.

When the route of the transmission area is calculated and the virtual link is activated, the result is logically equivalent to a point-to-point connection between the two endpoints. You can therefore configure the interface parameters on its physical interface and start the authentication function.

Unicast messages are transmitted between ABRs. A router that forwards a unicast message in the transmission area forwards it as an ordinary IP message. It can therefore only be understood as a logical link provided in the transmission area, over which protocol messages can be exchanged between two ABRs.

Mode: OSPFv3 Configuration Mode

Command: Area <area ID> virtual link <router ID>

Configures the transmission area and peer ID of the virtual link hello-interval <seconds> |

Configures the Hello interval of the virtual link dead-interval <seconds> |

Configures the neighbor failure time of the virtual link retransmit-interval <seconds> |

Configures the retransmission interval of the virtual link transmit-delay <seconds> |

Configures the interface delay of the virtual link.

Command: no area <area ID> virtual link <router ID> | hello-interval <seconds> | dead-interval <seconds> | retransmit-interval <seconds> | transmit-delay <seconds>]

Cancels the virtual link setting.

Parameter: area ID is the ID of the transmission area. It can use dotted decimal format a.b.c.d or integer format, with a value between 0 and 232-1. router ID is the ID of the peer router of the virtual link, in a.b.c.d format. Authentication and sending interface properties are optional; see the relevant command descriptions.

Default: The virtual link is not configured.

32.2.16 Configure Regional Routing Aggregation

Mode: OSPFv3 Configuration Mode

Command: Area <area ID> range <IPv6 prefix> [advertise | not advertise]

Configures the aggregation scope.

Command: no area <area ID> range <IPv6 prefix> [advertise | not advertise]

Cancels the aggregation.

Parameter: area ID is the area ID, which specifies the aggregation of routes in the area. You can use dotted decimal format a.b.c.d or integer format, with a value between 0 and 232-1. IPv6 prefix uses the prefix format X::X::X / m to represent the aggregation range. The optional parameters advertise and not advertise indicate whether to broadcast the aggregation range, that is, the IPv6 prefix. The original network routes are broadcast.

32.2.17 Configure Stub Area

Mode: OSPFv3 Configuration Mode

Command: Area <area ID> stub [no-summary]

Configures the router in the stub area.

Command: no area <area ID> stub [no summary]

Cancels the properties of the router in the stub area.

Command: Area <area ID> default cost <cost>

Configures the default cost of the ABR broadcast route connected to the stub area.

Command: no area <area ID> default cost

Restores the default cost to the default value.

Parameter: area ID is the area ID, indicating which area has the stub attribute. You can use dotted decimal format a.b.c.d or integer format, with a value between 0 and 232-1. no summary means that inter-domain routes are not injected into the stub area.

The first group of commands configures the router located in the stub area. The second group configures the ABR whose interface is connected to the stub area.

Default: The stub area is not configured.

32.2.18 Configure Default Weights for External Routes

The default value of the red metric command is not used when external weights are imported.

Mode: OSPFv3 Configuration Mode

Command: default metric <metric>

Configures the default weight used when external routes are introduced.

Command: no default metric [Metric]

Restores the default weight used when introducing external routes to the default value.

Parameter: metric value between 0 and 224-1

Default: The default weight is 1.

32.2.19 Display Information

Command: Show IPv6 OSPF [process ID]

Displays OSPFv3 process information.

Parameter: instance ID is the process number,

The value is between 0 - (216-1).

Command: Show IPv6 OSPF database <type>

Displays LSDB information.

Parameter: type refers to all types of LSA and summary information. See the command manual for details.

Command: Show IPv6 OSPF interface [if name]

Displays OSPFv3 interface information.

Parameter: if name is the agreed Layer 3 interface name.

Command: show ip ospf route

Displays the OSPF routing table.

Command: Show IPv6 OSPF virtual links

Displays OSPF virtual connection information.

Command: Show IPv6 OSPF neighbor [options]

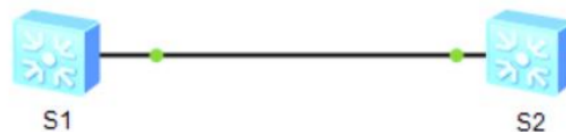
Displays OSPF neighbor information.

Parameter: options. See the command manual for details.

Mode: Privileged Mode

32.3 OSPFv3 Configuration Example

The two switches are connected as shown in the figure below.



The configuration of switch S1 is as follows:

```
Switch(config)#  
Switch(config)#int vlan1  
Switch(config-vlan1)#ipv6 address 3ffe:506::1/64  
Switch(config-vlan1)#exit  
Switch(config)#router ipv6 ospf  
Switch(config-router)#router-id 1.1.1.1  
Switch(config-router)#exit  
Switch(config)#int vlan1  
Switch(config-vlan1)#  
Switch(config-vlan1)#ipv6 router ospf area  
Switch(config-vlan1)#
```

The configuration of switch S2 is as follows:

```
Switch(config)#  
Switch(config)#int vlan1  
Switch(config-vlan1)#ipv6 address 3ffe:506::2/64  
Switch(config-vlan1)#exit  
Switch(config)#router ipv6 ospf  
Switch(config-router)# router-id 5.5.5.5  
Switch(config-router)#exit  
Switch(config)#int vlan1  
Switch(config-vlan1)#  
Switch(config-vlan1)#ipv6 router ospf area 0  
Switch(config-vlan1)#
```

Display the OSPFv3 neighbor information of the two switches.

```
Switch#show ipv6 ospf neighbor
```

OSPFv3 Process (*null*)

Neighbor ID Pri State Dead Time Interface Instance ID 5.5.5.5 1 Full/DR 00:00:39 vlan1 0

```
Switch#
```

Chapter 33 Configuring BGP

This chapter includes the following contents:

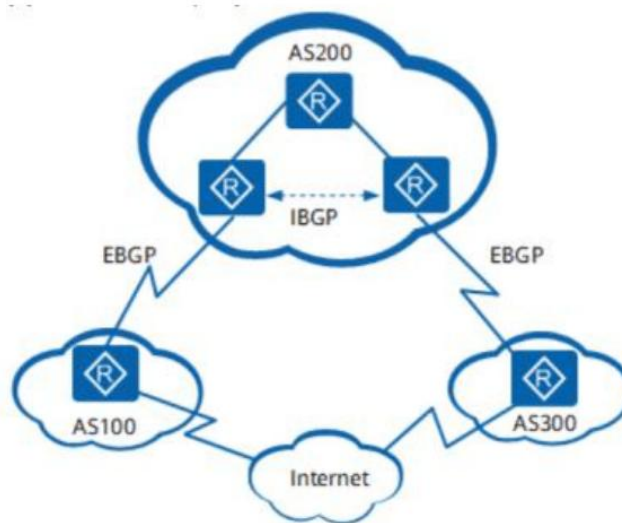
- PIM-SM Introduction
- PIM-SM Configuration
- PIM-SM Configuration Example

33.1 BGP Introduction

Border Gateway Protocol (BGP) is a distance vector routing protocol that provides routing reachability between autonomous systems (AS) and selects the best route. The three versions released earlier are BGP-1 (RFC1105), BGP-2 (RFC1163), and BGP-3 (RFC1267). BGP-4 (RFC1771) was used in 1994. After 2006, the version used in unicast IPv4 networks is BGP-4 (RFC4271), and the version used in other networks (such as IPv6) is MP-BGP (RFC4760).

According to its operation mode, BGP is divided into EBGP (external BGP) and IBGP (internal BGP).

The operation mode of BGP is shown in the figure below.



1 EBGP: BGP running between different ASs is called EBGP. To prevent a loop between ASs, when a BGP device receives a route sent by an EBGP peer, it discards the route that carries the local AS number.

2 IBGP: BGP running within the same AS is called IBGP. To prevent loops inside the AS, a BGP device does not announce a route learned from an IBGP peer to other IBGP peers, and it establishes a full connection with all IBGP peers. To solve the problem of too many IBGP peer connections, BGP defines the routing reflector and the BGP alliance.

There are five kinds of messages, six kinds of state machines, and five principles in the interactive process of establishing, updating, and deleting BGP peers.

BGP peers interact through the following five types of messages. Keepalive messages are sent periodically; the rest are triggered:

- Open message: used to establish a BGP peer connection.
- Update message: used to exchange routing information between peers.
- Notification message: used to interrupt a BGP connection.
- Keepalive message: used to maintain a BGP connection.
- Route refresh message: used to ask a peer to resend routing information after the routing policy changes. Only BGP devices that support the route refresh capability send and respond to this message.

There are six state machines in the BGP peer interaction process: idle, connect, active, open message sent (opensent), open message confirmed (openconfirm), and connection established (established). While BGP peers are being established, the three commonly visible states are idle, active, and established.

The BGP device adds the optimal route to the BGP routing table to form a BGP route. After a BGP device establishes a neighbor relationship with a peer, the following interaction principles apply:

- A BGP route obtained from an IBGP peer is published by the BGP device only to its EBGPs.
- A BGP route obtained from an EBGPs peer is published by the BGP device to all EBGPs and IBGPs.
- When there are multiple valid routes to the same destination address, the BGP device publishes only the optimal route to the peer.
- When a route is updated, the BGP device sends only the updated BGP route.
- A BGP device receives all routes sent by its peers.

BGP and IGP use different routing tables in the device. To allow different ASs to communicate with each other, BGP must interact with IGP: the BGP routing table and the IGP routing table are introduced into each other.

The BGP protocol itself does not find routes, so other routes must be introduced into the BGP routing table to achieve routing interworking between ASs. When an AS needs to publish routes to other ASs, the AS edge router introduces IGP routes into the BGP routing table. To plan the network better, when BGP introduces an IGP route it can use a route policy to filter the route and set route attributes, or set the MED value to guide the EBGPs peer in judging the route when traffic enters the AS.

BGP supports import and network when introducing routes:

- The import method imports the routes of rip, OSPF, Isis, and other protocols into the BGP routing table according to the protocol type. To ensure the effectiveness of the introduced IGP route, the import mode can also introduce static routes and direct routes.
- The network method introduces the existing routes in the IP routing table into the BGP routing table one by one. It is more accurate than the import method.

IGP introduces BGP routes.

When an AS needs to introduce routes from another AS, the AS edge router introduces BGP routes into the IGP routing table. To avoid the impact of a large number of BGP routes on the devices in the AS, when IGP introduces BGP routes you can use routing policies to filter the routes and set route attributes.

33.2 BGP Configuration

33.2.1 Start BGP and Enter BGP Configuration Mode

The BGP protocol is identified by an AS number. When you start the BGP protocol, you must specify which autonomous system is started.

Mode: Global Configuration Mode

Command: router BGP <as number>

Starts the BGP protocol process and enters its mode.

Command: no router BGP <as number>

Closes the BGP protocol and deletes its configuration.

Parameter: as number value: 1-4294967295.

Default: The BGP protocol is not run.

33.2.2 Inject Routing Information into the BGP Protocol

When BGP starts running, its routing information is empty. There are two ways to inject routing information into BGP: inject it manually with the network command, or inject it from IGP through the interaction with the IGP protocol. BGP publishes the injected route to its neighbors. This section describes the manual injection of routing information.

Mode: BGP Configuration Mode

Command: network <network number> mask <mask> [route map <map tag>] [backdoor]

This command configures the network information to be announced by this BGP speaker.

Command: no network <network number> mask <mask> [route map] [backdoor]

This command deletes the configured network information.

33.2.3 Configure BGP Timer

BGP uses the keepalive timer to maintain a valid connection with peers and the holdtime timer to judge whether peers are valid. By default, the keepalive timer is 60 seconds and the holdtime timer is 180 seconds. When a BGP connection is established between BGP speakers, both parties negotiate the holdtime, and the smaller holdtime value is selected. The keepalive timer value is the smaller of 1 / 3 of the negotiated holdtime and the configured keepalive value.

Mode: BGP Configuration Mode

Command: timers BGP <keepalive> <holetime>

This command adjusts the network timer of BGP.

Command: no timers BGP <keepalive> <holetime>

This command restores the default values.

33.2.4 Configure BGP and IGP Synchronization

For routing information that crosses this AS to another AS, the routing information can be published to the other AS only when all routers in this AS are known to have learned it. Otherwise, if a router in this AS (running the IGP protocol) has not learned the routing information, it may discard data messages that pass through this AS because it does not know the route, which causes the routing black hole phenomenon.

All routers in the AS must learn the routing information outside the AS. We call this the synchronization of BGP and IGP. The simple way to achieve synchronization is for the BGP speaker to redistribute all the routes learned by the BGP protocol to IGP, which ensures that the routers inside the AS can learn this routing information.

The synchronization mechanism of BGP can be canceled in two cases:

1. There is no routing information crossing this AS (generally, this AS is a terminal AS);
2. All routers in this AS run the BGP protocol, and all BGP speakers establish a full connection relationship (BGP speakers establish an adjacency relationship). Mode: BGP Configuration Mode

Commands: synchronization

This command starts the synchronization mechanism of BGP and IGP routing information.

Command: no synchronization

This command cancels the synchronization mechanism between BGP and IGP.

33.2.5 Configure the Interaction Between BGP and IGP

Inject the route information generated by the IGP protocol into BGP. By default, redistributing the default route is not allowed. To redistribute the default route, use the following command to control it.

Mode: BGP Configuration Mode

Command: redistribute <protocol type> [route map <map tag>]

This command redistributes the routing information of other routing protocols to BGP.

Command: no redistribute <protocol type> [route map <map tag>]

This command deletes the function and its parameter configuration.

33.2.6 Selection of BGP Optimal Path

The selection of the optimal route is a very important part of the BGP protocol. The selection process of the BGP routing protocol is described in detail below. If a routing table entry is invalid, it does not participate in the selection of the optimal route;

Otherwise, elect the routes with a high local_pref attribute value;

Otherwise, select the route generated by the BGP speaker;

The routes generated by this BGP speaker include those generated by the network command, the redistribute command, and the aggregate command.

Otherwise, select the route with the shortest as length;

Otherwise, select the route with the lowest origin attribute value;

Otherwise, select the route with the lowest MED value;

Otherwise, the EBGP path has a higher priority than the IBGP path and the AS alliance, and the IBGP path and the AS alliance have the same priority;

Otherwise, the route with the smallest IGP metric to the next hop is elected. Otherwise, among the routes from EBGP, the route received earlier is elected. Otherwise, the route announced by the BGP speaker with the smaller router ID is elected.

Otherwise, the route with the large cluster length is elected;

Otherwise, the route with the large neighbor address is elected.

Mode: BGP Configuration Mode

Command: BGP bestpath <as path ignore | compare confirmed aspath | compare routerid | don compare originator ID | Med [confirmed | missing as worst | remove rcv Med | remove send Med] | tie break on age>

This command sets the election of the BGP optimal route.

Command: no BGP bestpath <as path ignore | compare confirmed aspath | compare routerid | don compare originator ID | Med [confirmed | missing as worst | remove rcv Med | remove send Med] | tie break on age>

This command deletes the election of the BGP optimal route.

33.2.7 Configure BGP Routing Aggregation

BGP-4 supports CIDR, so you can create aggregation table entries to reduce the size of the BGP routing table. A BGP aggregation table entry is added to the BGP routing table only when there is a valid path within the aggregation range.

Mode: BGP Configuration Mode

Command: aggregate address <IP address> / <mask length> [as set | summary only]

This command sets the IPv4 / IPv6 aggregation routing table entry of BGP. Command: no aggregate address <IP address> / <mask length> [as set | summary only]

This command deletes the aggregation routing table entry.

33.2.8 Configure BGP Routing Reflector

To speed up the convergence of routing information, all BGP speakers in an AS usually establish a full connection relationship (BGP speakers establish an adjacency relationship two by two). When there are too many BGP speakers in the AS, this increases the resource overhead of the BGP speakers, increases the workload and complexity of configuration tasks for network administrators, and reduces the expansion performance of

the network. In this regard, two methods are proposed to reduce the number of IBGP peers in the AS: the routing reflector and the AS alliance. The routing reflector is a method to reduce the number of IBGP peer connections in an autonomous system. A BGP speaker is set as a routing reflector, which divides the IBGP peers in the autonomous system into two categories: client and non client.

The routing reflector is implemented in the AS, and its rules are as follows. Configure the routing reflector and specify its client. The routing reflector and its client form a group. Route Reflector

A connection relationship is established between it and the client.

The client of a routing reflector within a group should not establish a connection relationship with other BGP speakers outside the group.

In the AS, a complete connection relationship is established between the IBGP peers of non clients. Here, the IBGP of non clients

Peers include the following situations: between multiple routing reflectors in a group; between the routing reflectors in a group and, within the group,

BGP speakers that do not participate in the routing reflector function (usually these BGP speakers do not support the routing reflector function); and between the routing reflectors in the group and the routing reflectors in other groups.

The route reflector processes a received route according to the following rules:

A route update received from an EBGP speaker is sent to all clients and non clients;

A route update received from a client is sent to the other clients and all non clients;

Routing updates received from IBGP non clients are sent to all their clients.

Mode: BGP Configuration Mode

Command: neighbor [peer address | peer tag] route reflector client

This command configures the device as a routing reflector and specifies its client.

Command: no neighbor [peer address | peer tag] route reflector client

This command cancels the setting of the client.

33.2.9 Configure the BGP AS Federation

The federation is another way to reduce the number of IBGP peer connections in an autonomous system.

An autonomous system is divided into several sub autonomous systems, and these sub autonomous systems form an alliance by setting a unified alliance ID (that is, the alliance AS number). Outside the alliance, the whole alliance is still considered one AS, and only the AS number of the alliance is visible. Within the alliance, a complete IBGP peer connection is still established between the BGP speakers within a sub autonomous system, and an EBGP connection is established between the BGP speakers of different sub autonomous systems. Although an EBGP connection is established between the BGP speakers of the sub autonomous systems, path attribute information such as next_hop, MED, and local_pref remains unchanged when information is exchanged.

Mode: BGP Configuration Mode

Command: BGP confirmation identifier <as number>

This command configures the identifier of the AS federation.

Command: no BGP confirmation identifier

This command cancels the identifier of the AS federation.

33.2.10 Configure BGP Management Distance

The management distance indicates the reliability of a routing information source. It ranges from 1 to 255. The greater the value of the management distance, the lower the reliability.

BGP sets different management distances for the different sources of learned routing information. They are divided into external distance, internal distance, and local distance:

External distance: the management distance of routes learned from EBGp peers.

Internal distance: the management distance of routes learned from IBGP peers.

Local distance: routes learned from peers for which a better route is considered to be available from IGP

Management distance. These routes are usually indicated by the network backdoor command.

Mode: BGP Configuration Mode

Command: distance BGP <external distance> <internal distance> <local distance>

This command sets different management distances for different types of BGP routes.

Command: no distance BGP <external distance> <internal distance> <local distance>

This command restores the default management distance.

33.2.11 Configure BGP Routing Update Mechanism

The BGP routing update mechanism includes two parts: regular scan update and event triggered update. Regular scan update means that

BGP uses a timer to start the scanning mechanism regularly and update the routing table. Event triggered update means that when you configure BGP and a BGP configuration command changes or the next hop of a BGP route changes, BGP is triggered to start the scanning mechanism and update the routing table.

Mode: BGP Configuration Mode

Command: BGP scan time <time>

This command sets the time interval for regular scanning by the BGP protocol. Command: no BGP scan time [time]

This command restores the time interval for regular scanning by the BGP protocol to 60 seconds.

33.2.12 Configure BGP Local AS

The local AS function of BGP configures, for a specific peer, a local AS different from the real AS of BGP (router BGP as). This is similar to virtualizing a new AS with the peer device, so that when the real AS of the local BGP changes, the BGP configuration of the peer device does not need to be changed for the BGP connection to be established. This function is mainly used for

The migration and consolidation of ASs in large networks can ensure that the configuration of devices in other interconnected ASs is not affected. In the BGP protocol, when the local device establishes a BGP connection with a peer, it announces the local AS number to the peer device in the open message. The peer device checks whether the BGP AS announced by the connecting end is the same as the locally configured peer AS. If it is not, the peer device rejects the BGP connection. By default, the local AS in a BGP connection is the real AS of BGP. When you configure a local AS for a peer, the local device uses the configured local AS instead of the real AS when it establishes the BGP connection with the peer

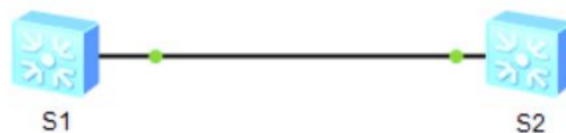
Mode: BGP Configuration Mode

Command: neighbor [peer address | peer tag] local as <as number>

Configures a local AS number for a BGP peer. The peer device can then use this local AS as its remoteas to establish a BGP connection with the device.

Command: no neighbor [peer address | peer tag] local as <as number> This command deletes the configured local AS.

33.3 BGP Configuration Example



The configuration of switch S1 is as follows:

```

Switch(config)#
Switch(config)#int vlan1
Switch(config-vlan1)#ip address 192.168.0.1/64
Switch(config-vlan1)#exit
Switch(config)#router bgp 1
Switch(config-router)#bgp router-id 1.1.1.1
Switch(config-router)#neighbor 192.168.0.2 remote-as 1
Switch(config-router)#neighbor 192.168.0.2 interface vlan1
Switch(config-router)# neighbor 192.168.0.2 next-hop-self
Switch(config-router)#exit
  
```

The configuration of switch S2 is as follows:

```
Switch(config)#
Switch(config)#int vlan1
Switch(config-vlan1)#ip address 192.168.0.2/64
Switch(config-vlan1)#exit
Switch(config)#router bgp 1
Switch(config-router)#bgp router-id 2.2.2.2
Switch(config-router)#neighbor 192.168.0.1 remote-as 1
Switch(config-router)#neighbor 192.168.0.1 interface vlan1
Switch(config-router)# neighbor 192.168.0.1 next-hop-self
Switch(config-router)#exit
```

Display the BGP neighbor information.

```
Switch#show bgp neighbors
```

BGP neighbor is 192.168.0.2, remote AS 1, local AS 1, internal link BGP version 4, remote router ID 2.2.2.2

BGP state = Established, up for 00:02:09

Last read 00:02:09, hold time is 90, keepalive interval is 30 seconds

Neighbor capabilities:

Route refresh: advertised and received (old and new)

Address family IPv4 Unicast: advertised and received

Received 7 messages, 0 notifications, 0 in queue

Sent 6 messages, 0 notifications, 0 in queue

Route refresh request: received 0, sent 0

Minimum time between advertisement runs is 5 seconds

For address family: IPv4 Unicast

BGP table version 2, neighbor version 2

Index 1, Offset 0, Mask 0x2

NEXT_HOP is always this router

Community attribute sent to this neighbor (both)

1 accepted prefixes

0 announced prefixes

Connections established 1; dropped 0

Local host: 192.168.0.1, Local port: 53557

Foreign host: 192.168.0.2, Foreign port: 179

Nexthop: 192.168.0.1

Nexthop global: fe80::82:44ff:fe13:4803

Nexthop local: ::

BGP connection: non shared network

Switch#

Chapter 34 Configuring VRRP

This chapter includes the following contents:

- VRRP Introduction
- VRRP Configuration
- VRRP Configuration Example

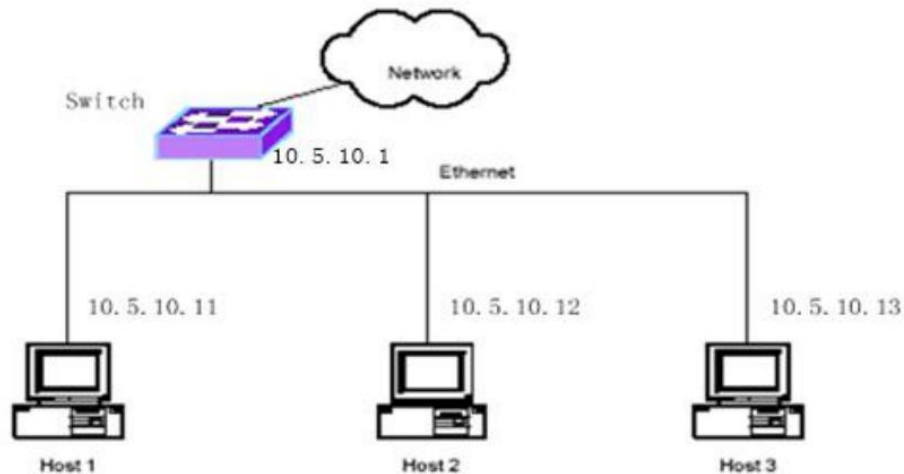
34.1 VRRP Introduction

VRRP is the abbreviation of Virtual Router Redundancy Protocol. It is an important three-layer reliability protocol used for redundant backup of the default gateway. This section describes the VRRP protocol in detail, and covers the following topics:

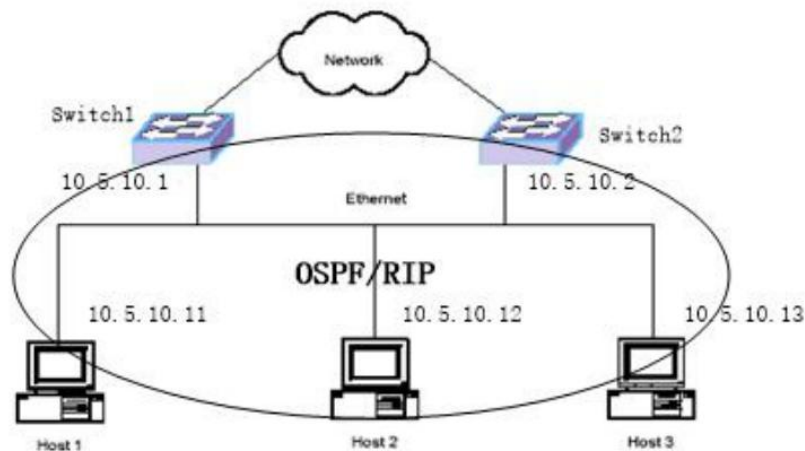
- VRRP Overview
- VRRP Terminology
- VRRP Protocol Interaction
- Election of the Virtual Master Router
- Status of the Virtual Router
- VRRP Tracing

34.1.1 VRRP Overview

The following figure shows a typical intranet networking scheme. One interface of the switch is connected to the external network, and the other interface is connected to the internal network. The IP address of the interface connected to the internal network is 10.5. 10.1, and hosts 1, 2, and 3 are all configured with IP addresses in the 10.5. 10.0/24 network segment. Hosts 1, 2, and 3 are configured with a default gateway whose next hop is the switch, and the IP address of the next hop is 10.5. 10.1. In this way, a message sent by a host whose destination IP address is not in the local network segment matches the default route and is sent to the switch, and the switch then forwards the message. The switch also forwards messages sent by the external network to the corresponding host, so the host can communicate with the external network.



In the above networking scheme, communication between a host and the external network can pass only through the single switch. When that switch fails, all the hosts are disconnected from the external network. To solve this problem, one solution is to expand one switch into two or more switches and run the dynamic routing protocol OSPF or RIP between the hosts and the switches, as shown in the following figure.

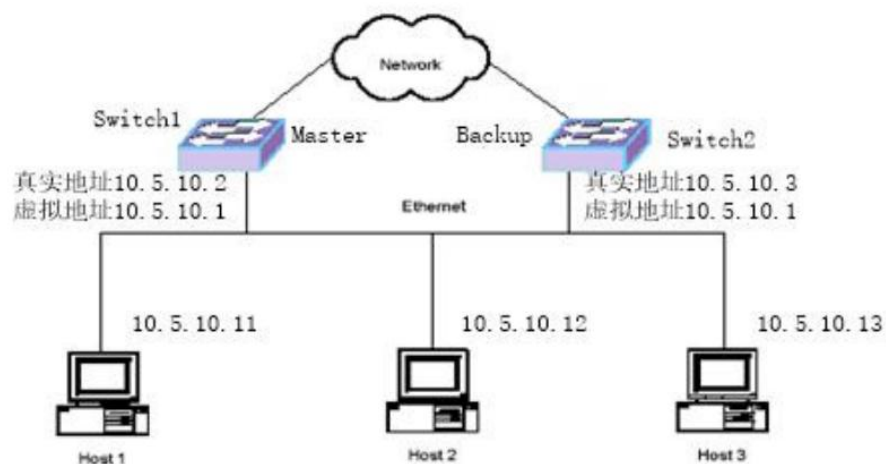


After a host runs the dynamic routing protocol, it can learn all the routes of the external network. When communicating with the external network, the host searches the route according to the destination IP address of the packet to get the next hop, and determines whether the packet is sent to switch1 or switch2. When one of the switches fails, the route in the host can be relearned in a very short time, and the next hop of the route points to the switch that has not failed, so communication between the host and the external network is not interrupted.

However, it is not practical to implement a dynamic routing protocol on the host. For hosts, running dynamic routing protocols is too much load. For networks, running dynamic routing protocols on hosts causes too much unnecessary data traffic on the network, and some hosts do not support dynamic routing protocols at all.

To solve the problem of a single point of failure, the VRRP protocol is the best choice. The VRRP protocol is proposed specifically for this problem. As shown in the figure below, Switch1 and Switch2 form a virtual router. The real IP addresses of the interfaces of the two switches are different, but there is a common virtual IP address 10.5.10.1, and the default gateway of the host is set to the virtual IP address 10.5.10.1. When Switch1 is the virtual master switch, communication between the host and the external network is forwarded through Switch1. When Switch1 fails, Switch2 takes over from Switch1 to become the virtual master switch, and communication between the host and the external network is forwarded through Switch2. With the VRRP protocol, the host only needs to set the default gateway and does not need to run other protocols. The load on the host is small, and only a few VRRP protocol flows need to be added on the network.

34.1.2 VRRP Terminology



The following terms are used frequently:

1) VRRP

Virtual Router Redundancy Protocol is a fault-tolerant protocol for default gateways that improves network reliability.

2) Virtual Router

A virtual router is an abstract object based on a subnet interface. It consists of a virtual router identifier (VRID) and an IP address, also known as the virtual IP address, which acts as the default gateway for hosts.

3) VRRP Router

VRRP routers are routers that run the VRRP protocol. A VRRP router can be added to a virtual router.

4) IP Address Owner

IP Address Owner. A VRRP router on which the virtual IP address of the virtual router is the same as the real IP address of the interface.

5) Virtual Router Master

The virtual master router forwards the Layer 3 data packets that pass through the virtual router and responds to ARP requests for the IP address of the virtual router. If a VRRP router is the IP address owner, it is always the virtual master.

6) Virtual Router Backup

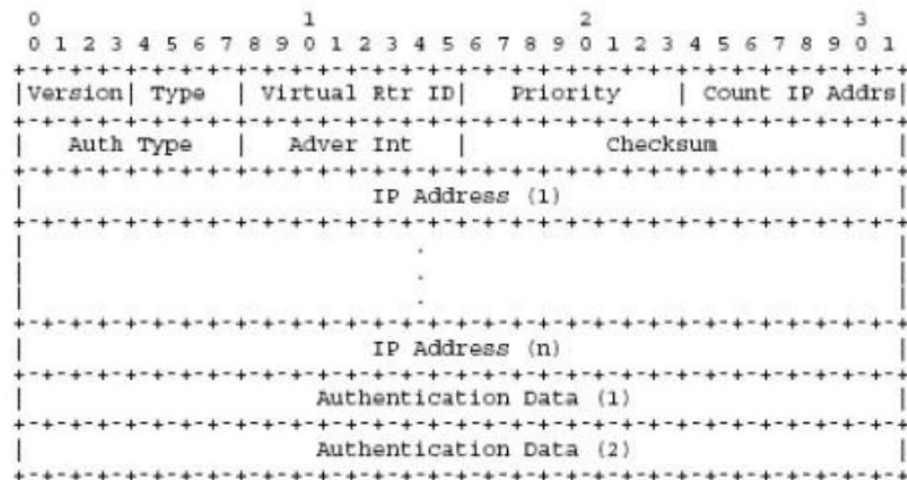
The virtual backup router does not forward Layer 3 packets and does not reply to ARP requests for the virtual IP address. It takes over the work of the virtual master router when the virtual master router fails.

To better understand these terms, note the following:

- A switch may include multiple interfaces, and the VRRP protocol may be enabled on multiple interface subnets.
- One virtual router can exist on one interface subnet.
- A VRID identifies a virtual router.

34.1.3 VRRP Protocol Interaction

The VRRP protocol packet is encapsulated in an IP packet. The VRRP message header is shown in the following figure:



1) MAC frame header field of the VRRP packet

Source MAC address: The virtual MAC address of the virtual router is 00-00-5e-00-01-{VRID}, where VRID is the virtual router identifier. For example, if the VRID of the virtual router is 1, the virtual MAC address is 00-00-5e-00-01-01.

Destination MAC address: The VRRP multicast MAC address, 01-00-5e-00-00-12.

2) IP header field of the VRRP packet

Source IP address: The primary IP address of the interface that sends the VRRP packet.

Destination IP address: The multicast IP address is 224.0.0.18. It cannot be used for Layer 3 forwarding.

TTL: 255, to prevent remote VRRP packet attacks.

Protocol: 112.

3) VRRP header field

Version: 2.

Type: the type of VRRP packet. Only one type is supported: 1 -- ADVERTISEMENT, the VRRP advertisement packet.

VRID: Identifies a virtual router.

Priority: The priority of the sending VRRP router for this virtual router.

Count IP Addrs: The number of virtual IP addresses. A virtual router can have multiple virtual IP addresses.

Auth Type: The authentication method between the VRRP routers in a virtual router.

Advertisement Interval: The interval between advertisements. The default is 1 second.

Checksum: The checksum, starting from the Version field of the VRRP header.

IP Address (es): One or more virtual IP addresses.

Authentication Data: The authentication data.

4) VRRP priority

Each VRRP router in a virtual router must be configured with a priority. The priority ranges from 0 to 255, where 0 and 255 are used for special purposes. The configurable priority ranges from 1 to 254. The default is 100. The higher the priority value, the higher the priority, and the more likely the router is to become the virtual master router.

In a virtual router, when a VRRP router is the IP address owner, its priority is 255.

When the virtual master router needs to notify the other backup routers that it is no longer the master router, it sends a VRRP packet with a priority of 0 to the other backup routers, which quickly triggers the other backup routers to become virtual master routers.

5) VRRP authentication

The VRRP protocol provides three authentication methods. In actual use, you can select one according to the security requirements of the network.

0 --- No Authentication

No authentication.

1 --- Simple Text Password

Simple password authentication.

2 --- IP Authentication Header

IP authentication header, which calculates a message digest by the HMAC-MD5 method.

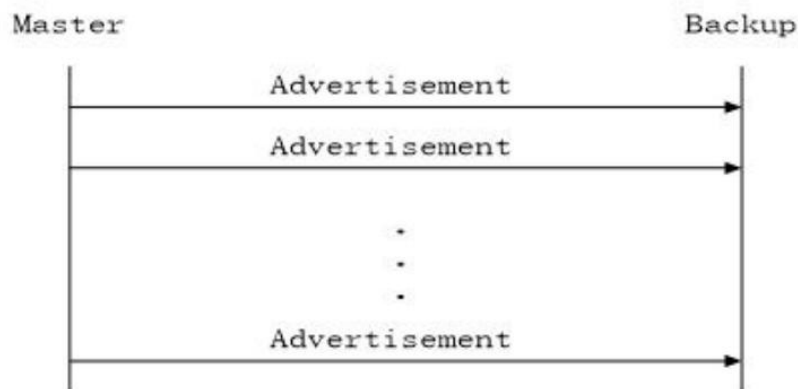
No authentication or simple password authentication can be used in a network with low security. HMAC authentication can be used in a network with high security.

For authentication methods 0 and 2, the Authentication Data field is filled with 0. For authentication method 1, the Authentication Data field is filled with the password. For authentication method 2, the message digest is filled in the IP Authentication Header field, that is, the AH field is added to the IP header.

6) VRRP packet interaction

The VRRP protocol has only one type of packet, the ADVERTISEMENT packet. In a virtual router, the virtual master router sends an Advertisement packet every Advertisement Interval (the default is 1 second). The virtual backup router determines whether a state migration is required according to the received VRRP advertisement packet. The protocol interaction between the master and the slave is shown in the figure below:

34.1.4 Election of the Virtual Master Router



The selection of the virtual master router in a virtual router is determined by the following factors:

- IP Address Owner

If a VRRP router is the IP address owner (its interface IP address is the same as the virtual IP address), it is the virtual master as long as the router is functioning properly.

- VRRP priority

The working VRRP router with the highest priority becomes the virtual master router. The configurable priority ranges from 1 to 254. The router that is the IP address owner has a priority of 255. When the virtual master router notifies the virtual backup router that it is no longer the master, a priority of 0 is given in the VRRP packet.

The size of the actual IP address of the interface. When the priorities are the same, the VRRP router with the larger actual interface IP address becomes the virtual master router.

An active-standby switchover occurs in a virtual router when:

1) When the virtual master router fails, an active/standby switchover occurs. In this case, there are two possibilities:

- If the virtual master router is still active, it sends a VRRP packet with a priority of 0. After receiving this packet, the backup router switches to the virtual master router if it does not receive a VRRP packet from the virtual master router within Skew_Time. In this case, the switching speed is relatively fast, and switching can be achieved within 1 second.
- If the virtual master router is not active, the virtual backup router switches to the virtual master router if it does not receive a VRRP packet from the virtual master router within the Master_Down_Interval.

$$\text{Master_Down_Interval} = (3 * \text{Advertisement_Interval}) + \text{Skew_Time}$$

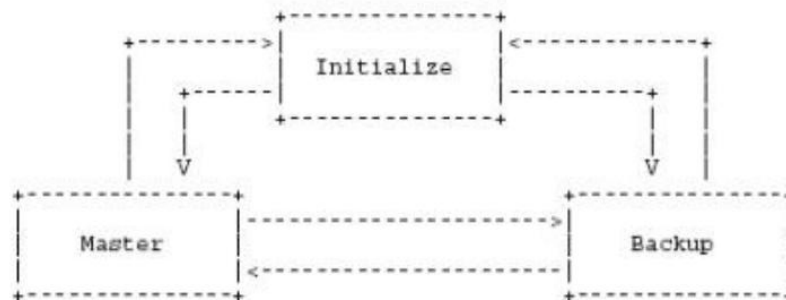
$$\text{Skew_Time} = ((256 - \text{Priority}) / 256)$$

2) When the virtual master router is not the IP address owner, but a router that is the IP address owner joins the network, that router becomes the virtual master router, and an active-standby switchover occurs.

3) When a VRRP router joins the network, if the priority of the router is higher than that of the virtual master router and the router is in preemption mode (the configuration variable Preempt_Mode is TRUE), the router becomes the virtual master router, and an active-standby switchover occurs.

34.1.5 Status of the Virtual Router

Each VRRP router in a virtual router runs a state machine. The migration of the state machine is shown in the following figure:



1) Initialize state

The Initialize state is the initial state of a virtual router, in which it waits for the Startup event. If a Startup event is received in this state, the process is as follows:

- If the priority of the VRRP router is 255 (that is, it is the IP address owner), the router becomes the virtual master router and moves to the Master state.
- Otherwise, the router becomes a virtual backup router and moves to the Backup state.

After the router moves to the Master state, the actions are as follows:

- Send a VRRP advertisement packet.
- Broadcast an ARP request containing the virtual IP address and the corresponding virtual MAC address.
- Set an Adver_Timer timer with a time interval of Advertisement_Interval.

After the router moves to the Backup state, the actions are as follows:

- Set a Master_Down_Timer timer with a time interval of Master_Down_Interval.

2) Backup state

The purpose of the Backup state is to monitor the availability and status of the virtual master router and to take over the work of the virtual master router at any time.

If a Shutdown event is received, cancel the Master_Down_Timer timer and return to the Initialize state.

If the Master_Down_Timer expires, the router becomes the virtual master router and moves to the Master state.

If a VRRP advertisement packet is received, the following conditions apply:

- If the priority field in the VRRP packet is 0, the Master_Down_Timer is set, and the timing interval is Skew_Time.
- Otherwise, if the preemption mode (Preempt_Mode) is FALSE or the priority in the VRRP packet $\geq$ the priority of the VRRP router, the Master_Down_Timer is reset, and the time interval is set to the Master_Down_Interval.
- Otherwise, discard the VRRP packet.

3) Master state

The VRRP router in the Master state forwards Layer 3 packets through the virtual router.

If a Shutdown event is received, cancel the Adver_Timer timer, send a VRRP advertisement packet with a priority of 0, and move to the Initialize state.

If the Adver_Timer timer expires, a VRRP advertisement packet is sent and the Adver_Timer timer is reset.

If a VRRP advertisement packet is received, the following conditions apply:

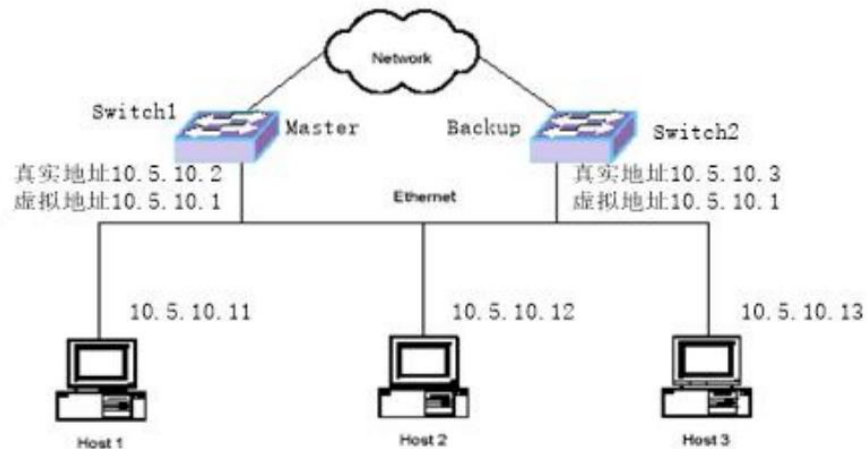
- If the priority in the packet is 0, send a VRRP advertisement packet and reset the Adver_Timer.
- Otherwise, if the priority in the packet is greater than that of the VRRP router, or the priorities are the same but the sending IP address is greater than the primary IP address of the VRRP router interface, cancel the Adver_Timer timer, set the Master_Down_Timer, and move to the Backup state.
- Otherwise, the VRRP advertisement packet is discarded.

34.1.6 VRRP Tracing

The VRRP protocol itself can only detect faults inside the virtual router, such as a LINK DOWN on the interface where the virtual router is located or a crash of the VRRP router. It cannot detect faults outside the virtual router. When a fault occurs outside the virtual router, the virtual router cannot use it to select the virtual master router, which interrupts network data. VRRP tracing solves this problem. The VRRP router traces specified external events. When an external failure occurs, the VRRP router changes its running priority and reselects the virtual master router so that network data is not interrupted.

As shown in the figure below, the external interface of the virtual master router Switch1 goes LINK DOWN. If the VRRP tracing function is not enabled, Switch1 cannot detect this external fault, Switch1 remains the virtual master router, and the host cannot access the external network. If the VRRP tracing function is enabled, Switch1 detects the external fault, modifies its running priority, and the virtual master router is reselected.

Switch1 becomes the virtual backup router and Switch2 becomes the virtual master router, so the host can continue to access the external network.



VRRP tracing includes interface tracing, route tracing, and PING tracing. Interface tracing means the VRRP router traces an interface outside the virtual router. If one of the traced interfaces goes LINK DOWN, an external fault has occurred. Route tracing means the VRRP router traces a route in its learned routing table. If the route does not exist, or exists but is not active, an external fault has occurred. PING tracing means the VRRP router continuously pings the traced device. If the device does not answer the PING within the specified time, an external fault has occurred. A virtual router can trace interfaces, routes, and PING at the same time. For each type of trace, multiple events can be traced. If any one of the traced events fails, the virtual router has an external failure. Only when all traced events are normal does the virtual router have no external failure.

34.2 VRRP Configuration

VRRP configuration includes the following:

- Create and delete a virtual router
- Configure the virtual IP address of the virtual router
- Configure parameters for the virtual router
- Configure VRRP tracing
- Start and shut down the virtual router
- View VRRP information

34.2.1 Create and Delete a Virtual Router

Virtual routers are created on subnet interfaces, and a VRID must be specified. No virtual router is created by default.

When a virtual router is no longer needed, you can delete it. If the virtual router is running, it is shut down first and then deleted. The commands to create and delete a virtual router are as follows:

Command	Description	CLI Mode
<code>router vrrp <vrid></code>	Create a virtual router and enter VRRP configuration mode. If the virtual router already exists, enter VRRP configuration mode directly. The argument is the VRID, which ranges from 1 to 255.	Configuration Mode
<code>no router vrrp [vrid]</code>	Delete a virtual router. The parameter is the VRID.	Configuration Mode

34.2.2 Configure the Virtual IP Address of the Virtual Router

A virtual IP address must be configured on the virtual router. In theory a virtual router can have one or more virtual IP addresses, but on this switch a virtual router supports only one virtual IP address. By default, no virtual IP address is configured on the switch.

The command to configure the virtual IP address of the virtual router is as follows:

Command	Description	CLI Mode
<code>virtual-ip <virtual-ip> <backup master></code>	Set the virtual IP address of the virtual router.	VRRP Configuration Mode
<code>no virtual-ip</code>	Delete the virtual IP address of the virtual router.	VRRP Configuration Mode

Notice:

- The virtual IP address of the virtual router can be configured only when the virtual router is shut down. It cannot be configured while the virtual router is running.
- The virtual IP address you set must be in the same network segment as the primary IP address of the interface; otherwise the configuration fails.

34.2.3 Configure Parameters for the Virtual Router

The parameters of the virtual router include priority, preemption mode, advertisement interval, authentication method, and authentication data. These parameters have default values, as shown in the following table:

Parameter	Default
Priority	100
Preemption Mode	TRUE
Advertisement Interval	1 Second
Authentication Method	Do Not Authenticate
Authentication Data	None

For a virtual router, the advertisement interval, authentication method, and authentication data must be configured identically, while the priority and preemption mode parameters may be configured identically.

Priority is divided into configuration priority and running priority. In most cases the running priority uses the configuration priority, but when the VRRP router is the IP address owner, the running priority is 255 and the configuration priority is not used.

For the authentication method, the switch currently implements only no authentication and simple password authentication. It does not implement the IP authentication header mode.

The commands for configuring the parameters of the virtual router are shown in the following table:

Command	Description	CLI Mode
<code>priority <priority-value></code>	Set the priority of the virtual router, from 1 to 254.	VRRP Configuration Mode
<code>preempt-mode {false true}</code>	Set the preemption mode of the virtual router. TRUE means preemption. FALSE means no preemption.	VRRP Configuration Mode
<code>advertisement-interval <interval></code>	Set the advertisement interval of the virtual router, from 1 to 255 seconds.	VRRP Configuration Mode
<code>authentication none</code>	Set the authentication method of the virtual router to no authentication.	VRRP Configuration Mode
<code>authentication simple-password <key></code>	Set the authentication method of the virtual router to simple password authentication, and set the authentication data, that is, the password.	VRRP Configuration Mode

Notice:

- The parameters of the virtual router can be configured only when the virtual router is shut down. They cannot be configured while the virtual router is running.

34.2.4 Configure VRRP Tracing

The switch currently implements only the VRRP interface tracking function. A VRRP router can track an interface, which can be a Layer 2 interface or an aggregation interface. By default, no tracked interface is configured on the switch.

You can configure a VRRP trace when the VRRP router is the IP address owner, but the VRRP trace does not actually take effect: the virtual master is not reselected even if the virtual router has an external failure. If you want to use the VRRP trace feature, do not configure the virtual router as the IP address owner.

VRRP tracing takes effect after you configure VRRP tracing, specify what to trace, and start the virtual router. When the VRRP router finds a tracked interface LINK DOWN, an external fault is considered to have occurred, and the running priority of the virtual router is set to the source priority minus priority-value. The virtual master router can then be reselected through the exchange of VRRP protocol packets. When the traced interfaces are all LINK UP, the failure is recovered and the running priority of the virtual router is reset to the configuration priority.

The commands for configuring VRRP tracing are shown in the following table:

Command	Description	CLI Mode
<code>circuit-failover <if-name> <priority-value></code>	Set the interfaces to be tracked by the virtual router.	VRRP Configuration Mode
<code>no circuit-failover</code>	Clear the tracked interfaces of the virtual router.	VRRP Configuration Mode

Notice:

VRRP tracing can be configured only when the virtual router is down. It cannot be configured while the virtual router is up.

34.2.5 Start and Shut Down the Virtual Router

After the virtual router is created and the virtual IP address and parameters are set, the virtual router is not yet running and is still in the Initialize state. Starting the virtual router starts the protocol, sends a Startup event to the protocol, and the state machine moves to the Master state or the Backup state. Shutting down the virtual router shuts down the protocol, sends a Shutdown event to the protocol, and returns the state to Initialize.

Make sure the virtual IP address is configured before you start the virtual router. If you need to modify the virtual IP address or the parameters while the virtual router is running, shut down the virtual router first, complete the configuration, and then start the virtual router again.

The commands to start and shut down the virtual router are as follows:

Command	Description	CLI Mode
<code>enable</code>	Start the virtual router.	VRRP Configuration Mode
<code>disable</code>	Shut down the virtual router.	VRRP Configuration Mode

34.2.6 View VRRP Information

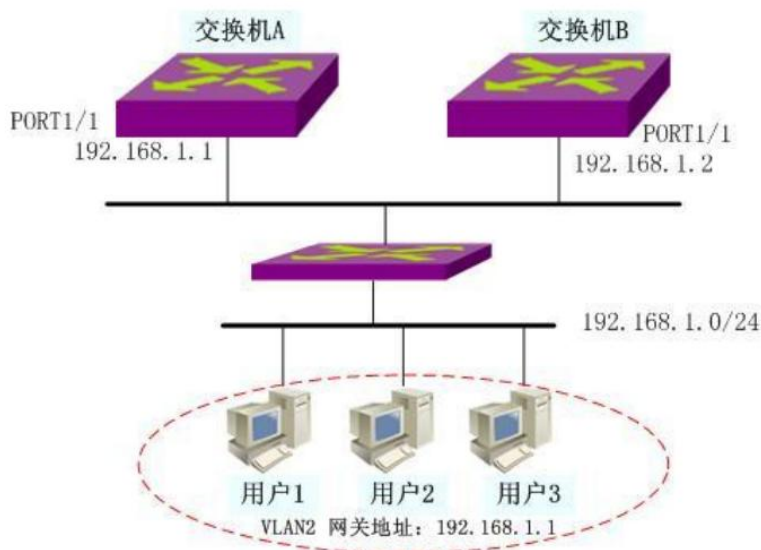
You can view the running status and configuration information of VRRP with a command. The command for viewing VRRP information is as follows:

Command	Description	CLI Mode
<code>show vrrp [vrid]</code>	If no parameters are entered, information about all virtual routers is displayed.	Normal Mode, Privileged Mode
<code>show running-config</code>	View the current configuration of the system. You can view the VRRP configuration.	Privileged Mode

34.3 VRRP Configuration Example

(1) Configuration

The VRRP function is enabled on the two switches to provide Layer 3 routing redundancy for users in the LAN and eliminate routing failure in the network. Switch 1 is set as the Master switch and Switch 2 is set as the Backup switch.



Configuration on Switch A:

```
Switch#configure terminal
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#exit
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#exit
Switch(config)#ip interface vlan 2
Switch(config)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.1/24
Switch(config-vlan2)#exit
Switch(config)#router vrrp 1
Switch(config-vrrp)# virtual-ip 192.168.1.1 master
Switch(config-vrrp)#enable
```

Configuration on Switch B:

```
Switch#configure terminal
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#exit
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#exit
Switch(config)#ip interface vlan 2
Switch(config)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.2/24
Switch(config-vlan2)#exit
Switch(config)#router vrrp 1
Switch(config-vrrp)# virtual-ip 192.168.1.1 backup
Switch(config-vrrp)#enable vrrp
```

(2) Verification

Use the following command to view information about VRRP:

```
show running-config
show vrrp
show vrrp 1
```

Chapter 35 Configuring VLLP

This chapter includes the following contents:

- Introduction to VLLP
- VLLP configuration
- VLLP configuration example

35.1 Introduction to VLLP

VLLP (VRRP Layer-2 Loop Protect Protocol) is a private protocol proposed to solve the problem of Layer 2 loops in VRRP applications. When two Layer 3 switches are used to provide redundant backup of the virtual route, a loop is formed between the Layer 3 switch and the Layer 2 switch. Through message exchange, the VLLP protocol informs each peer of the port status on the respective link, calculates the loop, and sets the state of a specific port to block so as to cut off the loop. A port set to the block state is reset to the forward state when the timer expires. The VLLP protocol monitors and maintains the loop state of the port by notifying port state changes and timer timeouts in time, and ensures that the loop is cut off in time. The VLLP protocol uses a query-reply approach to gather port status information. When two Layer 3 switches running the VLLP protocol start up, one is automatically elected to be in the sender state and the other in the receiver state. The sender sends the query message regularly, and the receiver receives the query message and sends back the response message. When the receiver has a port status change, it actively sends a link status change message to notify the sender to make a corresponding change. The VLLP protocol works with the VRRP protocol to start and maintain the state of the relevant ports in the VLAN. The VLLP protocol only needs to run on the VRRP switch; the Layer 2 switch does not need to run any loop protection protocol.

Basic concepts of the VLLP protocol:

VLLP device: a VLLP protocol entity running on a VLAN.

VLLP port: a port that participates in VLLP protocol exchange in the VLAN corresponding to the VLLP device.

VLLP device status: a VLLP device has two states, sender and receiver.

a. Sender: the VLLP device in the sender state actively sends a VLLP query message at regular intervals.

b. Receiver: the VLLP device in the receiver state responds to the query message and actively sends a VLLP link state change message when the link state changes.

VLLP port status: a VLLP port has three STP states: disable, block, and forward.

Master port: when the VLLP device is in the sender state, one VLLP port is selected as the master port. A VLLP sender has only one master port. A VLLP port with a high priority and a mapping relationship can preempt the master port role.

VLLP port mapping relationship: a VLLP port that exchanges VLLP protocol messages with the peer switch.

VLLP device receiver election principles:

a. The device with the higher priority becomes the receiver. b. If the priorities are the same, the device with the largest MAC address becomes the receiver.

Master port election principles:

a. The port must be link up. b. The VLLP port must have a mapping relationship.

If no VLLP port satisfies the conditions, there is no master port.

If multiple VLLP ports meet the conditions, the one with the higher priority is selected as the master port. If the priorities are the same, the one with the earliest mapping relationship is selected as the master port.

VLLP port status determination principles:

If the VLLP device is the sender, the state of the master port must be forward.

If the VLLP device is the sender, the link state is link up, and the VLLP port has no mapping relationship, the port state is forward.

If the VLLP device is the sender, the link state is link up, and the VLLP port has a mapping relationship, the port state is block.

If the VLLP device is the receiver and the link state is link up, the VLLP port state is forward.

A VLLP port whose link is down has the disable state.

The VLLP protocol message is encapsulated in a MAC frame.

**Destination
MAC Address (6
bytes)**

Source MAC
Address (6
bytes)

0x8100		Prio	Vlan ID (12 bits)		Ethernet Type (2 bytes)
Version	Type	Port1 (2 bytes)			
Priority	Query Inter	Main port			
Reserved (4 bytes)					
Port2		Link state		STP state	

There are three types of VLLP protocol messages:

Link state query message LQ

Link state response message LA

Link state change message LC

Value range of each field in the message format:

Des MAC: fixed to 00:09:ca:ff:ff

Src MAC: MAC of the VLAN sending the VLLP protocol message

Ethernet Type: fixed as 0x268e

Version: currently 1

Type: LQ is 1; LA is 2; LC is 3

Port 1: index value of the port sending the VLLP protocol message

Priority: VLLP device priority, value 1 ~ 255

Query Interval: VLLP sender query timer interval, 5 seconds by default

Main port: index value of the main port of the VLLP sender, present only in the LQ message; Reserved: the reserved field is zero

Port 2: in the LC message, the index value of the port whose status changed; in the LQ and LA messages, port2 is the same as port1

Link state: the link state of port2, where link up is 1 and link down is 2

STP state: the STP state of port2, where disable is 1, block is 2, and forward is 3.

VLLP protocol principle:

A VLLP device is configured in one VLAN and the VLLP protocol is started, and a VLLP device is also configured in the corresponding VLAN of the peer switch and the VLLP protocol is started. The VLLP protocol entities (VLLP devices) running on the VLAN then form a sender and receiver pair. When the protocol starts, both sides are senders and send LQ messages to each other. When a VLLP device receives an LQ message, it elects a receiver according to the priority carried in the message and the MAC address of the peer. The winner becomes the receiver and responds to the sender's LQ message instead of sending LQ messages. When the receiver's timer times out and no LQ message has been received, the receiver returns to the sender state and starts sending LQ messages.

In a VLAN where the VLLP protocol is enabled, a port that participates in VLLP protocol message exchange must be configured as a VLLP port. A VLLP port can be any valid Layer 2 port, including a trunk group, but a trunk member cannot be configured as a VLLP port. VLLP protocol messages are sent and received through VLLP ports. A VLLP port and the corresponding VLLP port configured in the VLAN of the peer switch that runs the VLLP protocol form a pair of mapping relationships. The mapping relationship with the peer port is determined by sending a query message and receiving a response message or a change message. Loops that may exist in the network are calculated from the mapping relationship and its link state, and loops in the topology are stopped by maintaining the STP state of the ports according to the VLLP port state determination.

Multiple VLLP ports can be enabled within a VLAN that has the VLLP protocol enabled, and they may or may not be physically linked to the peer switch. When a port belongs to multiple VLANs, the same VLLP port can also appear in multiple VLLP devices. The VLLP protocol dynamically collects information about link state

changes of the VLLP port and the STP state of the peer VLLP port, so it can calculate the loop in time and effectively prevent loops in the network.

When multiple VLANs have the same port configuration but the VLLP protocol must be started on several of them, each Layer 2 port has to send and receive multiple VLLP protocol messages running on different VLANs, which burdens the switch. The concept of the attached VLAN addresses this. VLANs with identical port configurations run the VLLP protocol on only one VLAN, the primary VLAN, while the remaining VLANs are added as secondary VLANs on the instance of the primary VLAN. The result of the loop calculation performed by the VLLP protocol on the primary VLAN is written to the port state within the instance. When you configure the secondary VLAN, make sure that all VLLP ports of the primary VLAN are in the secondary VLAN and that all ports of the secondary VLAN are in the primary VLAN. When the attached VLAN is configured and a Layer 2 port is added to it, if the port is also in the primary VLAN, the primary VLAN manages the port state uniformly; if the port is not in the primary VLAN, its state cannot be managed and an alarm message is displayed.

35.2 VLLP Configuration

After the VLLP protocol is started, you can configure the related attributes and create ports. The related commands are in VLLP configuration mode.

VLLP configuration includes:

- Create a VLLP device on a Layer 3 interface
- Enable the VLLP device
- Create a VLLP port on a Layer 2 interface
- Configure the VLLP device priority
- Configure the VLLP device query timer interval
- Configure an affiliated VLAN
- Configure the VLLP port priority.

35.2.1 Create a VLLP Device on a Layer 3 Interface

Mode: Global Configuration Mode

Command: `router vllp <if-name>`

Create a VLLP device and enter VLLP configuration mode.

Command: `no router vllp <if-name>` Remove the VLLP device

Parameter: if-name is the agreed Layer 3 interface name (for example: `vlan1 vlan2 ...`)

Default: the VLLP protocol is not started

35.2.2 Enable the VLLP Device

Mode: VLLP Configuration Mode

Command: `vllp enable` Enable the VLLP device

Command: vllp disable Disable the VLLP device

Default: the VLLP device is not started after creation

35.2.3 Create a VLLP Port on a Layer 2 Interface

Mode: VLLP Configuration Mode

Command: vllp port <if-name> Create a VLLP port

Command: no vllp port <if-name> Delete the VLLP port

Parameter: if-name is the agreed Layer 2 interface name (for example: ge1/1 trunk1 ...)

Default: the VLLP protocol is not applied to the Layer 2 port. The VLLP protocol cannot be applied when the Layer 2 interface is a trunk member.

35.2.4 Configure the VLLP Device Priority

Mode: VLLP Configuration Mode

Command: vllp priority <priority> Configure the VLLP device priority

Command: no vllp priority [priority]

Restore the VLLP device priority to the default.

Parameter: the priority value is between 1 and 255. The VLLP device uses the priority to elect a receiver.

Default: 100

35.2.5 Configure the VLLP Device Query Timer Interval

Mode: VLLP Configuration Mode

Command: vllp query-interval <interval>

Configure the local query timer interval.

Command: no vllp query-interval [interval]

Restore the query timer interval to the default.

Parameter: interval is between 1 and 255. The configured value takes effect when the VLLP device is the sender or when it migrates back to the sender.

Default: 5 seconds

35.2.6 Configure an Affiliated VLAN

Mode: VLLP Configuration Mode

Command: vllp dependency <if-name>

```
configure an affiliated VLAN
```

Command: no vllp dependency <if-name> Remove the attached VLAN

Parameter: if-name is the agreed Layer 3 interface name (for example: vlan1 vlan2 ...)

Default: no secondary VLAN is configured

35.2.7 Configure the VLLP Port Priority

Mode: VLLP Configuration Mode

Command: vllp port <if-name> priority <priority>

Configure the VLLP port priority.

Command: no vllp port <if-name> priority [priority]

Restore the VLLP port priority to the default.

Parameter: if-name is the agreed Layer 2 interface name (for example: ge1/1 trunk1 ...); the priority value is between 1 and 255. The VLLP sender uses the priority to elect the primary port.

Default: 100

35.2.8 Display Information

Mode: Normal or Privileged

Command: show vllp

Display the list of VLLP devices in the VLLP protocol.

Command: show vllp <if-name>

Display detailed information about a VLLP device.

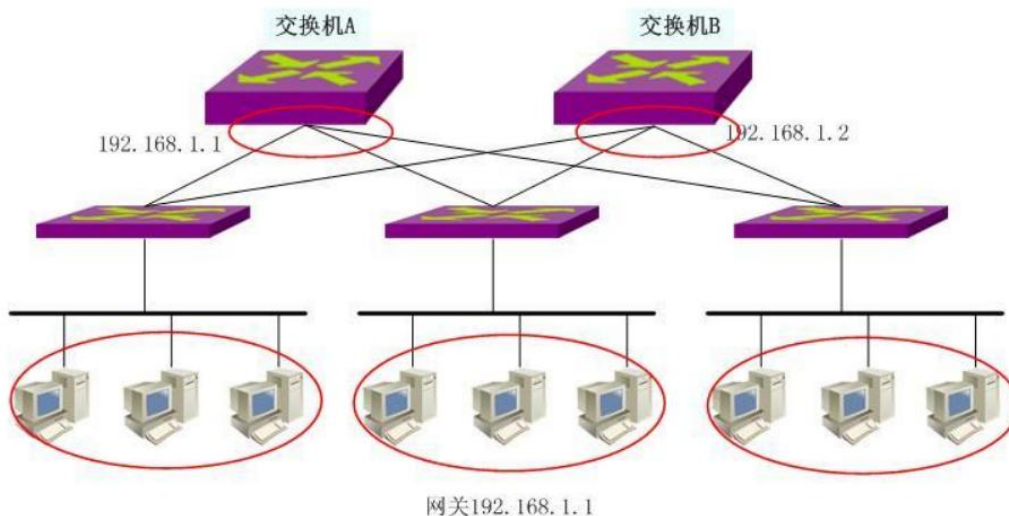
Parameter: if-name is the agreed Layer 3 interface name (for example: vlan1 vlan2 ...)

Command: show vllp map

Display the mapping relationship of each VLLP port in the VLLP protocol.

35.3 VLLP Configuration Example

(1) Configuration



Configuration on Switch A:

```
Switch#configure terminal
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#exit
Switch(config)#ip interface vlan 2
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport access vlan 2
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#switchport access vlan 2
Switch(config-ge1/3)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.1/24
Switch(config-vlan2)#exit
Switch(config)#router vrrp 1
Switch(config-vrrp)# virtual-ip 192.168.1.1 master
Switch(config-vrrp)#enable
Switch(config-vrrp)#exit
Switch(config)#router vllp vlan2
Switch(config-vllp)#vlllp port ge1/1
Switch(config-vllp)#vlllp port ge1/2
Switch(config-vllp)#vlllp port ge1/3
Switch(config-vllp)#vlllp enable
```

Configuration on Switch B:

```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 2
Switch(config-vlan)#ip interface vlan 2
Switch(config-vlan)#exit
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport access vlan 2
Switch(config-ge1/1)#interface ge1/2
Switch(config-ge1/2)#switchport access vlan 2
Switch(config-ge1/2)#interface ge1/3
Switch(config-ge1/3)#switchport access vlan 2
Switch(config-ge1/3)#interface vlan2
Switch(config-vlan2)#ip address 192.168.1.2/24
Switch(config-vlan2)#exit
Switch(config)#router vrrp 1
Switch(config-vrrp)# virtual-ip 192.168.1.1 backup
Switch(config-vrrp)#virtual-interface vlan1
Switch(config-vrrp)#enable
Switch(config-vrrp)#exit
Switch(config)#router vllp vlan2
Switch(config-vllp)#vlllp port ge1/1
Switch(config-vllp)#vlllp port ge1/2
Switch(config-vllp)#vlllp port ge1/3
Switch(config-vllp)# enable
```

(2) Verification

Use the following command to view VLLP information:

```
show vllp
show vllp <if-name>
show vllp map
```

Chapter 36 Configuring Policy Routing

This chapter includes the following contents:

- Introduction to Policy Routing
- Policy routing configuration
- Policy routing configuration example

36.1 Introduction to Policy Routing

Policy-based routing is a routing mechanism based on the user's policy. Instead of simply looking up the routing table by the destination address of the IP packet to forward it, policy routing selects the route flexibly based on attributes of the packet, such as the destination address, the source address, and other information. This enriches the routing knowledge of the router.

36.2 Policy Routing Configuration

The configuration includes the following:

- Create a new policy route
- Insert a policy route
- Delete a policy route
- Move a policy route
- View policy routing information

36.2.1 Create a New Policy Route

Use the following command in global configuration mode to create a new policy route:

```
policy route <ID> <SIP> <DIP> <next-hop>
```

ID: the ID of the newly created policy routing rule, ranging from 1 to 100.

SIP, DIP: there are three input methods for the source and destination IP addresses:

1) A.B.C.D wildcard can control the IP addresses of one network segment.

2) any is equivalent to A.B.C.D 255.255.255.255

3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0 wildcard: determines which bits must be matched. '0' means a match is required, and '1' means no match is required.

next-hop: the address of the next-hop host, in the format A.B.C.D.

36.2.2 Insert a Policy Route

Use the following command in global configuration mode to insert a new policy route:

```
policy route insert <ID> <SIP> <DIP> <next-hop> before <EXIST_ID>
```

ID: the ID of the newly inserted policy rule, ranging from 1 to 100.

SIP, DIP: there are three input methods for the source and destination IP addresses:

- 1) A.B.C.D wildcard can control the IP addresses of one network segment.
- 2) any is equivalent to A.B.C.D 255.255.255.255
- 3) host A.B.C.D is equivalent to A.B.C.D 0.0.0.0 wildcard: determines which bits must be matched. '0' means a match is required, and '1' means no match is required.

next-hop: the address of the next-hop host, in the format A.B.C.D.

EXIST_ID: the rule to insert before. Range: 1-100.

36.2.3 Delete a Policy Route

Use the following command in global configuration mode to delete a policy route:

```
no policy route <ID>
```

ID: the ID of the policy rule to delete, in the range 1-100.

36.2.4 Move a Policy Route

Use the following command in global configuration mode to move a policy route to a destination:

```
policy route move <ID> (before|after) <TO_ID>
```

ID: the rule to move.

(Before after): whether to move before or after the target rule.

36.2.5 View Policy Routing Information

Run the following command in normal or privileged mode to view all policy routing information:

```
show policy route
```

36.3 Policy Routing Configuration Example

Configure the source IP address as 192.168.3.100 and the gateway as 192.168.0.20.

```
Switch#configure terminal
Switch#(config)#policy route 1 host 192.168.3.100 any 192.168.0.20
```

Configure a gateway with a destination IP address of 192.168.10.100 and route to 192.168.2.1.

```
Switch#configure terminal
Switch#(config)#policy route 2 any host 192.168.10.100 192.168.2.1
```

Configure the source IP address as 192.168.3.100, the destination IP address as 10.10.10.100, and the gateway as 192.168.5.1.

```
Switch#configure terminal
Switch#(config)#policy route 3 host 192.168.3.100 host 10.10.10.100 192.168.5.1
```

Chapter 37 Configuring the System Log

This chapter includes the following contents:

- System log introduction
- Syslog configuration

37.1 System Log Introduction

The system log module is an important part of the switch. It records the operation of the whole system, abnormal behavior, and user operations, and helps administrators understand and monitor the work of the system in time. The system log module manages all the log information generated by the running modules of the system, and collects, classifies, stores, and displays that log information.

The log system also includes an important debugging function. With debugging, system logs help administrators and other technicians monitor the operation of the network and debug and diagnose network faults. You can select the content to debug and then locate and solve a fault in the device or network by observing the log information that debugging outputs.

This section includes the following topics:

- Format of log information
- Log storage
- Log display
- Debugging tool

37.1.1 Format of Log Information

The format of the log information is as follows:

Timestamp, priority, module name, log content

A space separates the timestamp and the priority, a colon and a space separate the priority and the module name, and a colon and a space separate the module name and the log content.

An example of the log information format is as follows:

```
2006/05/20 13:56:34 Warning: MSTP: Port up notification received for port ge1/2
```

In this log message, the timestamp is 2006/05/20 13:56:34, the priority is Warning, the module name is MSTP, and the log content is Port up notification received for port ge1/2.

1) Timestamp

Format of the timestamp: year/month/day hour:minute:second.

The hour uses a 24-hour clock, from 0 to 23.

The timestamp records the time when the log information was generated, using the system time of the switch. The system time is set before the switch leaves the factory and can be modified by the administrator. The system time keeps running after the device is powered off.

2) Priority

The priority records the importance of the log information. Log information is divided into four levels by importance. In order of priority from high to low, they are Critical, Warning, Informational, and Debugging. The priorities are described in the following table:

Priority	Description
Critical	Serious error
Warning	General errors, warnings, and very important tips
Informational	Important tips, general tips, diagnostic information
Debugging	Debugging information

3) Module name

The module name records the module that generated the log information. The following table lists some of the main modules that generate log information:

Module Name	Description
CLI	Command line interface module
MSTP	Multi-instance Spanning Tree Protocol module
VLAN	VLAN function module
ARP	ARP protocol module
IP	IP protocol module
ICMP	ICMP protocol module
UDP	UDP protocol module
TCP	TCP protocol module

4) Log content

The log content is a phrase or sentence that states the main point of the log information. Administrators can learn what happened to the system by reading the log content.

37.1.2 Log Storage

Logs are generally stored in three ways:

- The log is stored in memory.
- The log is stored in NVM.
- The log is stored on a server.

Memory holds four log tables, one per log priority. Each table stores the log information of one priority, so logs are divided into four categories by priority and each category is stored in a separate log table. Each log table has 1K entries and can store 1K log messages. When a log table is full, the next log overwrites the oldest log information. This storage method has a problem: when the system restarts, the log information is gone, so the administrator cannot see the log information from the time the system crashed and cannot locate the problem.

Important log information, such as log information with a priority of Critical or Warning, can be stored in the NVM of the system. The log information in NVM is retained after the system restarts, which makes it easier for the administrator to locate the problem when the system crashes. However, this storage method has a problem: because of the capacity limitation of NVM, the number of log entries stored in NVM is very limited.

A better way is to store the log information on a server, which can be done with the SYSLOG protocol. The log information is sent to the server in real time, and the server saves it and displays it on an interface. This storage method makes log information easy to view and has a huge capacity, so a large amount of log information can be stored on the server.

Currently, the system supports storing log information only in memory. It does not support storing log information in NVM or on a server.

37.1.3 Log Display

There are two ways to display the log: manual display and real-time display. Manual display means you display the log information by entering commands. Real-time display means that when log information is generated, it is output directly to the terminal so you can see it immediately.

In manual display mode, you can view all the log information. The log information is displayed with the most recent entry first, so you see the latest operating status of the switch first.

In real-time display mode, you must turn on the terminal real-time display switch. If the switch is on, the generated log information is written to the log table and also output to the terminal. If the switch is off, the log information is not displayed on the terminal in real time. Currently, the system can output log information in real time only to the Console terminal. It does not support outputting log information to a Telnet terminal.

37.1.4 Debugging Tool

Debugging is a diagnostic tool for devices and networks. It tracks the sending and receiving of data packets by the system and its modules, and the state machine changes of the modules, so the administrator can understand and monitor the running process of the system and its modules. If the network or a device is abnormal, you can track it with the debugging tool.

The debugging tool provides many switches that administrators can control to track what they are interested in. When the device or network is abnormal, the administrator can turn on the debugging switch related to the abnormality and find the problem by tracing the execution process of the system and the module.

When a debugging switch is turned on, the system generates the relevant log information and writes it to the corresponding log table. In general, the priority of log information generated by debugging is Informational. When the terminal real-time display switch is turned on, the log information is output to the terminal in real time. When the debugging switch is turned off, the system does not generate the relevant log information.

37.2 Syslog Configuration

System log configuration includes the following:

- Configure the terminal real-time display switch
- View log information
- Configure the debugging switch
- View debugging information

37.2.1 Configure the Terminal Real-Time Display Switch

By default, the terminal real-time display switch is turned off: the log information generated by the system is written to the log table but is not displayed on the terminal in real time. Some log information in the system is not restricted by this switch, and that log information is always output to the Console terminal in real time.

Currently, the switch can display log information in real time only on the Console terminal, not on a Telnet terminal.

When you use the write command to store the current configuration of the system in the configuration file, the configuration of the terminal real-time display switch is not stored in the configuration file of the system. When the system restarts, this configuration is lost and must be set again.

The commands for configuring the terminal real-time display switch are shown in the following table:

Command	Description	CLI Mode
log stdout	Turn on the terminal real-time display switch	Configuration Mode
no log stdout	Turn off the terminal real-time display switch	Configuration Mode

37.2.2 Set Log Level

The commands for setting the log level are shown in the following table:

Command	Description	CLI Mode
<code>log trap <[alerts critical debugging emergencies errors informational notifications warnings]></code>	Set the log level	Configuration Mode

37.2.3 View Log Information

The commands for viewing log information are shown in the following table:

Command	Description	CLI Mode
<code>show log</code>	Show all log information	Configuration Mode

37.2.4 Configure the Debugging Switch

The system provides many debugging switches across multiple modules. Only representative commands for each module are listed here. For the complete command format, see the command manual.

When you use the write command to store the current configuration of the system in the configuration file, the configuration of the debugging switch is not stored in the configuration file of the system. When the system restarts, this configuration is lost and must be set again.

Representative commands for configuring the debugging switch are as follows:

Command	Description	CLI Mode
<code>debug ip ...</code>	Turn on the debugging switches related to the system sending and receiving IP packets	Privileged Mode
<code>no debug ip ...</code>	Turn off the debugging switches related to the system sending and receiving IP packets	Privileged Mode
<code>debug ip icmp ...</code>	Turn on the debugging switches related to the system sending and receiving ICMP packets	Privileged Mode
<code>no debug ip icmp ...</code>	Turn off the debugging switches related to the system sending and receiving ICMP packets	Privileged Mode
<code>debug ip arp ...</code>	Turn on the debugging switches related to the system sending and receiving ARP packets	Privileged Mode
<code>no debug ip arp ...</code>	Turn off the debugging switches related to the system sending and receiving ARP packets	Privileged Mode
<code>debug ip udp ...</code>	Turn on the debugging switches related to the system sending and receiving UDP packets	Privileged Mode
<code>no debug ip udp ...</code>	Turn off the debugging switches related to the system sending and receiving UDP packets	Privileged Mode

Command	Description	CLI Mode
<code>debug ip tcp ...</code>	Turn on the debugging switches related to the system sending and receiving TCP packets	Privileged Mode
<code>no debug ip tcp ...</code>	Turn off the debugging switches related to the system sending and receiving TCP packets	Privileged Mode
<code>debug mstp ...</code>	Turn on the debugging switch related to MSTP protocol diagnosis	Privileged Mode
<code>no debug mstp ...</code>	Turn off the debugging switch related to MSTP protocol diagnosis	Privileged Mode
<code>debug igmp snooping ...</code>	Turn on the debugging switch related to IGMP SNOOPING protocol diagnosis	Privileged Mode
<code>no debug igmp snooping ...</code>	Turn off the debugging switch related to IGMP SNOOPING protocol diagnosis	Privileged Mode
<code>debug dhcp snooping ...</code>	Turn on the debugging switch related to DHCP SNOOPIN protocol diagnosis	Privileged Mode
<code>no debug dhcp snooping ...</code>	Turn off the debugging switch related to DHCP SNOOPIN protocol diagnosis	Privileged Mode
<code>no debug all</code>	Turn off all debugging switches on the system	Privileged Mode

37.2.5 View Debugging Information

The command to view debugging information is as follows:

Command	Description	CLI Mode
	View the debugging switch configuration. If no parameter is entered, the debugging switch configuration of all modules is shown. If one parameter is entered, only that module's debugging switch configuration is shown. If the parameter is IP, the debugging switch configuration of the IP, ICMP, ARP, UDP, and TCP modules is shown.	Normal Mode, Privileged Mode

37.3 Configuring SYSLOG

SYSLOG includes the following:

- Introduction to SYSLOG
- SYSLOG configuration
- SYSLOG configuration example

37.3.1 Introduction to SYSLOG

SYSLOG is a standard protocol for managing device log information. Its simple design has made it widely used. The SYSLOG system is divided into three parts.

The first part defines each sub-module, so the log information generated by different modules can be distinguished, and defines the log information levels, so the operating status of the equipment can be observed. All log information of the device is collected according to this convention.

The second part is the configuration file, which defines how the collected log information is processed. It can be saved locally, sent to a designated server on the network, or distributed to designated logged-in users. The configuration file determines how the log information generated by the device is saved.

The third part sends the SYSLOG protocol message in the message format defined by the RFC. In our switch system, the working convention of the whole SYSLOG is the system log module.

The first part of the SYSLOG protocol is handled by each functional sub-module in the switch, which sends its own log information of each level to the system log module. The syslog module maintains four levels of log tables.

In the second part of the SYSLOG protocol, the system log module distributes the log information uniformly. First, the log information is displayed on the serial port terminal in real time or manually through the terminal display switch. Second, four levels of log tables are stored in memory. Third, high-level log information is stored in NVM so that important log records are not lost in case of power failure. Fourth, the log is sent to a remote server through a SYSLOG message for storage, collection, and collation. The SYSLOG submodule in the system log module implements only the third part of the function, transmitting the system log to the server.

37.3.2 SYSLOG Configuration

The SYSLOG configuration commands include:

- Open the syslog server address
- Close the syslog server address
- Open the syslog protocol

Command	Description	CLI Mode
<code>syslog open <server-ip></code>	Open the syslog server address; the parameter server-ip is the server IP address.	Configuration Mode
<code>syslog close</code>	Close the syslog server address	Configuration Mode
<code>log syslog</code>	Open the syslog protocol	Configuration Mode

37.3.3 SYSLOG Configuration Example

(1) Configuration

Configure the IP address of the syslog server as 192.168.0.201, and configure the switch as follows:

```
Switch#configure terminal
Switch(config)#syslog open 192.168.0.201
Switch(config)# log syslog
```

(2) Verification

```
Switch#show syslog
```

Syslog is opened!

server ip address: 192.168.0.201 udp destination port: 514 severity level: debugging local device name: switch

Chapter 38 IPv6 Basic Configuration

The switch supports basic IPv6 functions, including IPv6 Layer 2 forwarding and IPv6 ND functions. This chapter describes how to configure IPv6 and includes the following:

- Introduction to IPv6
- Configure basic functions of IPv6
- Configure the IPv6 Neighbor Discovery Protocol
- IPv6 display and maintenance

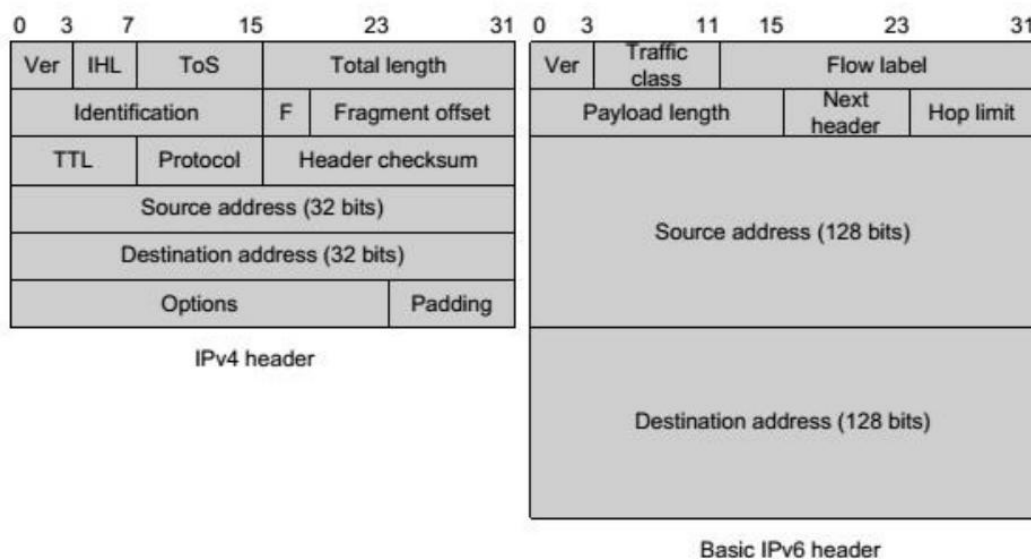
38.1 Introduction to IPv6

IPv6 (Internet Protocol Version 6) is the second generation standard protocol of the network layer, and is also called IPng (IP Next Generation). It is a set of specifications designed by the IETF (Internet Engineering Task Force) and an upgraded version of IPv4. The most significant difference between IPv6 and IPv4 is that the length of the IP address has increased from 32 bits to 128 bits.

38.1.1 IPv6 Protocol Features

(1) Simplified header format

The length of the IPv6 basic header is reduced by removing some fields of the IPv4 header or moving them into the extension header. IPv6 uses a fixed-length basic packet header, which simplifies the processing of IPv6 packets by forwarding devices and improves forwarding efficiency. Although the IPv6 address is four times as long as the IPv4 address, the IPv6 basic header is only 40 bytes long, which is twice as long as the IPv4 header (excluding the options field).



Comparison between the IPv4 header and the IPv6 basic header

(2) Sufficient address space

The source address and destination address of IPv6 are 128 bits (16 bytes) long. This provides more than 3.4×10^{38} possible addresses, which fully meets the needs of multi-level address division as well as the address allocation of public networks and private networks within institutions.

(3) Hierarchical address structure

The IPv6 address space uses a hierarchical address structure, which supports fast routing lookup and, with route aggregation, effectively reduces the system resources occupied by the IPv6 routing table.

(4) Automatic address configuration

To simplify host configuration, IPv6 supports both stateful and stateless address configuration:

1) Stateful address configuration means obtaining an IPv6 address and related information from a server, such as a DHCP server.

2) Stateless address configuration means that the host automatically configures the IPv6 address and related information from its own link layer address and the prefix information issued by the router.

The host can also form a link-local address from its own link layer address and the default prefix (FE80::/10) to communicate with other hosts on the link.

(5) Built-in security

IPv6 uses IPSec as its standard extension header, which provides end-to-end security. This also provides a standard for solving network security problems and improves interoperability between different IPv6 applications.

(6) QoS support

The Flow Label field of the IPv6 packet header identifies the flow, allowing the device to identify a packet in a given flow and give it special processing.

(7) Enhanced neighbor discovery mechanism

The IPv6 neighbor discovery protocol is implemented through a set of ICMPv6 (Internet Control Message Protocol for IPv6) messages, and manages the exchange of information between neighbor nodes, that is, nodes on the same link. It replaces ARP (Address Resolution Protocol), ICMPv4 router discovery, and ICMPv4 redirect messages, and provides a range of other features.

(8) Flexible extension headers

IPv6 removes the option field of the IPv4 header and introduces a variety of extension headers. This improves processing efficiency, greatly enhances the flexibility of IPv6, and gives the IP protocol good scalability. The option field in the IPv4 header is at most 40 bytes, while the size of an IPv6 extension header is limited only by the size of the IPv6 packet.

38.1.2 IPv6 Address Introduction

(1) IPv6 address representation

IPv6 addresses are written as a series of 16-bit hexadecimal numbers separated by colons (:). Each IPv6 address is divided into 8 groups. The 16 bits of each group are represented by 4 hexadecimal digits, and the groups are separated by colons, for example, 2001:0000:130F:0000:0000:09C0:876a:130B.

To simplify the representation of an IPv6 address, the "0" in the IPv6 address can be handled as follows:

- 1) The leading "0" in each group can be omitted, so the address above can be written as 2001:0:130F:0:0:9C0:876A:130B.
- 2) If the address contains two or more consecutive groups of 0, they can be replaced by a double colon ":", so the address above can be written as 2001:0:130F:9C0:876A:130B.

Notice:

The double colon ":" can be used only once in an IPv6 address. Otherwise, when the device changes the ";;" back to zeros to recover the 128-bit address, the number of zeros it represents cannot be determined.

An IPv6 address consists of two parts: an address prefix and an interface identifier. The address prefix corresponds to the network number field of the IPv4 address, and the interface identifier corresponds to the host number part of the IPv4 address.

The address prefix is written as: IPv6 address/prefix length. The IPv6 address is in any of the forms listed above, and the prefix length is a decimal number indicating how many of the leftmost bits of the IPv6 address are the address prefix.

(2) IPv6 address classification

IPv6 has three main types of addresses: unicast address, multicast address, and anycast address.

- 1) Unicast address: uniquely identifies an interface, similar to the unicast address of IPv4. Datagrams sent to a unicast address are delivered to the interface identified by that address.
- 2) Multicast address: identifies a group of interfaces, usually belonging to different nodes, similar to the multicast address of IPv4. Datagrams sent to a multicast address are delivered to all interfaces identified by that address.
- 3) Anycast address: identifies a group of interfaces, usually belonging to different nodes. A datagram sent to an anycast address is delivered to the interface in that group that is closest to the source node, as measured by the routing protocol in use.

IPv6 has no broadcast address; the function of the broadcast address is provided by the multicast address.

The IPv6 address type is specified by the first few bits of the address, called the format prefix. The correspondence between the main address types and the format prefixes is shown in Table 1-1.

	Address Type	Format Prefix (Binary)	IPv6 Prefix Identification
Unicast address	No address specified	00...0 (128 bits)	::/128
	Loopback address	00...1 (128 bits)	::1/128
	Link-local address	1111111010	FE80::/10
	Site-local address	1111111011	FEC0::/10
	Global unicast address	Other forms	-
Multicast address		11111111	FF00::/8
Anycast address		Taken from the unicast address space, using the format of a unicast address	

Table 1-1 Correspondence between Address Type and Format Prefix

(3) Types of unicast address

There are many types of IPv6 unicast addresses, including global unicast addresses, link-local addresses, and site-local addresses.

1) The global unicast address is equivalent to the IPv4 public network address and is provided to the network service provider. This type of address allows aggregation of routing prefixes, which limits the number of global routing table entries.

2) Link-local addresses are used for communication between nodes on a link-local in Neighbor Discovery Protocol and Stateless Autoconfiguration. Datagrams that use link-local addresses as source or destination addresses are not forwarded to other links.

3) Site-local addresses are similar to private addresses in IPv4. Datagrams that use a site-local address as the source or destination address are not forwarded to other sites outside the site, which is equivalent to a private network.

4) Loopback address: the unicast address 0:0:0:0 (abbreviated as: 1) is called the loopback address and cannot be assigned to any physical interface. Its function is the same as the loopback address in IPv4: a node uses it to send IPv6 messages to itself.

5) Unspecified address: the address ":" is called the unspecified address and cannot be assigned to any node. Before a node obtains a valid IPv6 address, it can put this address in the source address field of the IPv6 messages it sends, but it cannot be used as the destination address in an IPv6 message.

(4) Multicast address

The multicast addresses shown in Table 1-2 are reserved for special purposes.

Address	Application
FF01::1	All-node multicast address, node-local scope
FF02::1	All-node multicast address, link-local scope
FF01::2	All-router multicast address, node-local scope
FF02::2	All-router multicast address, link-local scope
FF05::2	All-router multicast address, site-local scope

Table 1-2 List of Reserved IPv6 Multicast Addresses

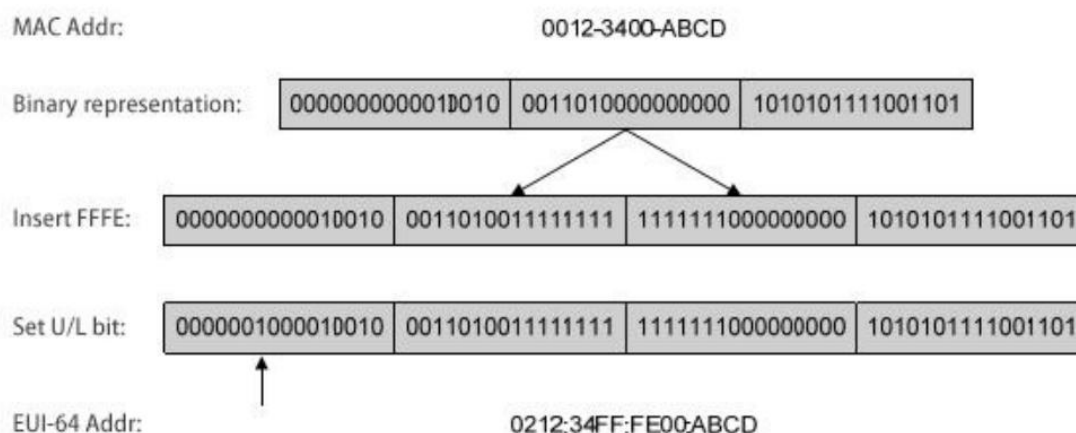
There is another type of multicast address: the solicited-node address. This address is used mainly to obtain the link layer address of a neighbor node on the same link and to perform duplicate address detection. Each unicast or anycast IPv6 address has a corresponding solicited-node address. The format is:

FF02:0:0:0:1:FFXX:XXXX

Here, FF02:0:0:0:1:FF is a fixed 104-bit format, and XX:XXXX is the last 24 bits of the unicast or anycast IPv6 address.

(5) Interface identifier in IEEE EUI-64 format

The interface identifier in an IPv6 unicast address identifies a unique interface on a link. At present, IPv6 unicast addresses generally require a 64-bit interface identifier. The interface identifier in IEEE EUI-64 format is derived from the link layer address (MAC address) of the interface. Because the interface identifier in an IPv6 address is 64 bits and a MAC address is 48 bits, the hexadecimal number FFFE (111111111111110) is inserted in the middle of the MAC address, after the 24th bit from the high order end. To ensure that the interface identifier derived from the MAC address is unique, the Universal/Local (U/L) bit, which is bit 7 from the high order end, is also set to "1". The resulting set of numbers is used as the interface identifier in EUI-64 format.



MAC address to EUI-64 format interface identifier

38.1.3 Introduction to the IPv6 Neighbor Discovery Protocol

The IPv6 Neighbor Discovery Protocol uses five types of ICMPv6 messages to implement the following functions: address resolution, neighbor reachability verification, duplicate address detection, router discovery/prefix discovery, address autoconfiguration, and redirection.

The types and functions of the ICMPv6 messages used by the Neighbor Discovery Protocol are shown in Table 1-3.

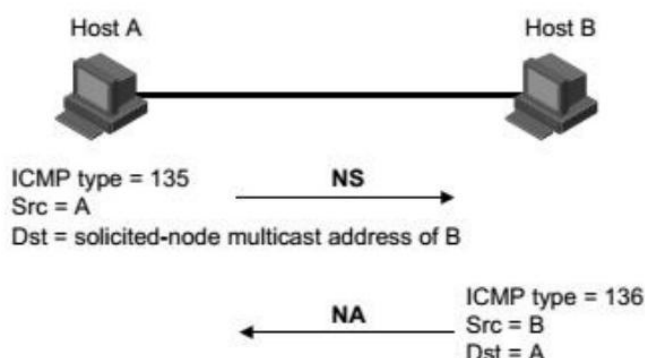
ICMPv6 Message	Type No.	Action
Neighbor Solicitation	135	Obtains the link layer address of a neighbor
		Verifies whether a neighbor is reachable
		Performs duplicate address detection
Neighbor Advertisement	136	Responds to an NS message
		When the link layer changes, the node sends an NA message to notify neighbor nodes of the change
Router Solicitation	133	After startup, the node sends an RS message to the router to request a prefix and other configuration information for automatic configuration of the node
Router Advertisement	134	Responds to an RS message
		If RA messages are not suppressed, the router issues them periodically. An RA message includes the prefix information option and some flag bits
Redirect	137	When certain conditions are met, the default gateway sends a redirect message to the source host so that the host reselects a correct next-hop address for subsequent messages.

Table 1-3 ICMPv6 Message Types and Functions Used by the Neighbor Discovery Protocol

The Neighbor Discovery Protocol provides the following main functions:

(1) Address resolution

Address resolution obtains the link layer address of a neighbor node on the same link, the same function that ARP provides in IPv4. It uses the neighbor solicitation message NS and the neighbor advertisement message NA. In the following figure, node A wants to obtain the link layer address of node B.



Address Resolution Diagram

- 1) Node A sends an NS message by multicast. The source address of the NS message is the IPv6 address of node A's interface, the destination address is the solicited-node multicast address of node B, and the message content includes the link layer address of node A.
- 2) After receiving the NS message, node B checks whether the destination address of the message is the solicited-node multicast address that corresponds to its own IPv6 address. If it is, node B learns the link layer address of node A and returns an NA message by unicast that carries its own link layer address.
- 3) Node A obtains the link layer address of node B from the received NA message.

(2) Verify whether the neighbor is reachable

After the link layer address of a neighbor node is obtained, you can verify whether the neighbor node is reachable by using the neighbor solicitation message NS and the neighbor advertisement message NA.

- 1) The node sends an NS message whose destination address is the IPv6 address of the neighbor node.
- 2) If a confirmation message from the neighbor node is received, the neighbor is considered reachable. Otherwise, the neighbor is considered unreachable.

(3) Duplicate address detection

When a node acquires an IPv6 address, it uses duplicate address detection to determine whether the address is already used by another node, similar to gratuitous ARP in IPv4. Duplicate address detection uses NS and NA messages, as shown in Figure 1-4.

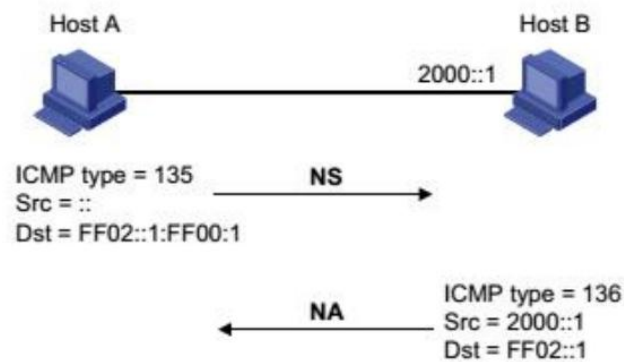


Figure 1-4 Duplicate Address Detection Diagram

1) Node A sends an NS message. The source address of the NS message is the unspecified address ::, the destination address is the solicited-node multicast address that corresponds to the IPv6 address to be detected, and the message content includes the IPv6 address to be detected. 2) If node B is already using this IPv6 address, it returns an NA message that contains its own IPv6 address.

3) When node A receives the NA message from node B, it knows that the IPv6 address is already in use. Otherwise, the address is not in use and node A can use it.

(4) Router discovery/prefix discovery and address auto-configuration

Router discovery/prefix discovery means that the node obtains the neighbor router and the prefix of the network where it is located, as well as other configuration parameters, from the received RA message.

Stateless address autoconfiguration means that the node automatically configures its IPv6 address from the information obtained by router discovery/prefix discovery. Router discovery/prefix discovery uses router solicitation messages RS and router advertisement messages RA. The process is as follows:

1) When the node starts, it sends an RS message to the router to request the prefix and other configuration information needed to configure the node.

2) The router returns an RA message that includes the prefix information option. The router also issues RA messages periodically.

3) The node uses the address prefix and other configuration parameters in the RA message returned by the router to automatically configure the IPv6 address and other information of the interface.

- The prefix information option includes the address prefix as well as the preferred lifetime and valid lifetime of that prefix. After receiving a periodic RA message, the node updates the preferred lifetime and valid lifetime of the prefix according to the message.
- During the valid lifetime, the automatically generated address can be used normally. After the valid lifetime expires, the automatically generated address is deleted.

(5) Redirection function

When a host starts, its routing table may contain only a default route to the default gateway. When certain conditions are met, the default gateway sends an ICMPv6 redirect message to the source host to notify it to

select a better next hop for subsequent packets. This function is the same as that of the IPv4 ICMP redirect message.

- The device sends an ICMPv6 redirect message that redirects the host when the following conditions are met:
- The interface that receives the data message and the interface that forwards it are the same interface.
- The selected route was not itself created or modified by an ICMPv6 redirect message.
- The selected route is not the default route.
- The forwarded IPv6 data packet does not contain a routing extension header.

38.1.4 IPv6 PMTU Discovery

The links along the transmission path from the source to the destination may have different MTUs. In IPv6, when a packet is longer than the MTU of the link, the packet is fragmented at the source. This reduces the processing load on the intermediate forwarding devices and makes rational use of network resources.

The purpose of the PMTU (Path MTU) discovery mechanism is to find the minimum MTU on the path from the source to the destination. Figure 1-5 shows how PMTU discovery works.

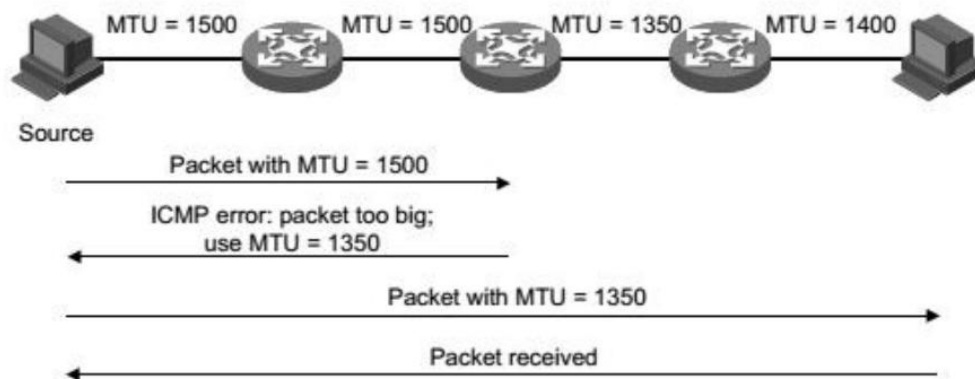


Figure 1-5 Working Process of PMTU Discovery

- 1) The source host fragments the message using its own MTU and then sends the message to the destination host.
- 2) When an intermediate forwarding device receives the message and forwards it, if the MTU supported by the forwarding interface is smaller than the length of the message, the device discards the message and returns an ICMPv6 error message to the source. The error message contains the MTU of the interface that failed to forward.
- 3) After receiving the error message, the source host uses the MTU carried in the message to fragment the message again and sends it.
- 4) This process repeats until the destination host receives the message, which determines the minimum MTU on the path of the message from the source to the destination.

38.1.5 Protocol Specification

The protocol specifications related to the IPv6 foundation are:

- RFC 1881: IPv6 Address Allocation Management
- RFC 1887: An Architecture for IPv6 Unicast Address Allocation
- RFC 1981: Path MTU Discovery for IP version 6
- RFC 2375: IPv6 Multicast Address Assignments
- RFC 2460: Internet Protocol, Version 6 (IPv6) Specification.
- RFC 2461: Neighbor Discovery for IP Version 6 (IPv6)
- RFC 2462: IPv6 Stateless Address Autoconfiguration
- RFC 2463: Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification
- RFC 2464: Transmission of IPv6 Packets over Ethernet Networks
- RFC 2526: Reserved IPv6 Subnet Anycast Addresses
- RFC 3307: Allocation Guidelines for IPv6 Multicast Addresses
- RFC 3513: Internet Protocol Version 6 (IPv6) Addressing Architecture
- RFC 3596: DNS Extensions to Support IP Version 6

38.2 Introduction to Basic IPv6 Configuration Tasks

Configure basic functions of IPv6

Configure the IPv6 Neighbor Discovery Protocol

Configure PMTU discovery

Configure ICMPv6 message sending

38.3 Configure the Basic Functions of IPv6

38.3.1 Configure the IPv6 Unicast Address

IPv6 global unicast addresses are obtained by manual assignment.

IPv6 link-local addresses are obtained in two ways:

- Automatic generation: when the VLAN port is UP, the device automatically generates a link-local address for the interface from the link-local address prefix (FE80::/10) and the link layer address of the interface.
- Manual designation: you configure the IPv6 link-local address manually.

Command	Description	CLI Mode
<code>ipv6 address <ipv6-address>/<prefix-length></code>	Specify the IPv6 address manually. By default, link-local addresses are generated automatically under layer 3 interfaces from the MAC address of the VLAN interface.	Configuration Mode

38.4 Configure the IPv6 Neighbor Discovery Protocol

38.4.1 Configure the Relevant Parameters of the RA Message

You can configure whether an interface sends RA messages and the interval at which it sends them. You can also configure the parameters carried in the RA message to notify the host. When the host receives the RA message, it can use these parameters to act accordingly. The parameters that can be configured in RA messages and their meanings are shown in Table 1-4.

Table 1-4 Parameters and Descriptions in the RA Message

Parameter	Description
Cur Hop Limit	When the host sends an IPv6 message, it uses this parameter value to fill the Hop Limit field in the IPv6 message header. The device also uses this value for the Hop Limit field in its response messages.
Prefix Information	After a host on the same link receives the prefix information issued by the device, it can perform operations such as stateless automatic configuration.
Managed address configuration flag bit (M flag)	Determines whether a host acquires an IPv6 address by stateful autoconfiguration. If this bit is set to 1, the host obtains an IPv6 address through stateful autoconfiguration, such as from a DHCP server. Otherwise, the host obtains an IPv6 address through stateless autoconfiguration, which generates an IPv6 address from its link layer address and the prefix information published by the router.
Other configuration flag bit (O flag)	Determines whether a host uses stateful autoconfiguration to obtain information other than an IPv6 address. If the Other Configuration flag bit is set to 1, the host obtains the additional information through stateful autoconfiguration, for example from a DHCP server. Otherwise, it obtains the additional information through stateless autoconfiguration.
Router Lifetime	The time for which the router that issued the RA message is used as the default router for the host. From the router lifetime value in the received RA message, the host determines whether to use the router that issued the RA message as its default router.
Retrans Timer	After the device sends an NS message, it resends the NS message if it receives no response within the specified time interval.
Reachable Time	After neighbor reachability detection confirms that a neighbor is reachable, the device treats the neighbor as reachable for the configured reachable time. Once that time is exceeded, the device reconfirms whether the neighbor is reachable before sending a message to it.
Link MTU	The MTU option is used in the RA message to ensure that all nodes on the link use the same MTU value, mainly when the nodes may not know the MTU of the link. Other Neighbor Discovery messages must silently ignore this option.

Configure the hop count limit

Command: `ipv6 nd cur-hop-limit value`

View Mode: VLAN Interface Mode

Default configuration: by default, the router is limited to 64 hops

Unsuppress RA message publishing

Command: `ipv6 nd send-ra`

View Mode: VLAN Interface Mode

Default configuration: by default, publishing of RA messages is suppressed

Configure the maximum and minimum time intervals for RA message publishing

Command: `ipv6 nd max-ra-interval value`

View Mode: VLAN Interface Mode

Default configuration: by default, the maximum interval for RA message publishing is 600 seconds

Command: `ipv6 nd min-ra-interval value`

View Mode: VLAN Interface Mode

Default configuration: by default, the minimum time interval for RA message publishing is 198 seconds

Notice:

- When RA messages are issued periodically, the interval between two consecutive messages falls between the minimum time interval and the maximum time interval; a value is selected at random from that range as the interval for issuing the RA message.
- The configured minimum time interval must be less than or equal to 0.75 times the maximum time interval.

Configure the prefix information in the RA message

Command: `ipv6 nd prefix X:X::X:X/M (valid-lifetime preferred-lifetime (off-link | no-autoconfig))`

View Mode: VLAN Interface Mode

Default configuration: by default, the prefix information in the RA message is not configured. In that case, the IPv6 address of the interface that sends the RA message is used as the prefix information in the RA message.

Set the managed address configuration flag bit

Command: `ipv6 nd managed-config-flag`

View Mode: VLAN Interface Mode

Default configuration: by default, the managed address flag bit is 0, that is, the host obtains the IPv6 address through stateless autoconfiguration.

Set the other configuration flag bit

Command: `ipv6 nd other-config-flag`

View Mode: VLAN Interface Mode

Default configuration: by default, the other configuration flag bit is 0, that is, the host obtains other information through stateless automatic configuration.

Configure the router lifetime in the RA message

Command: `ipv6 nd ra-lifetime value`

View Mode: VLAN Interface Mode

Default configuration: by default, the router lifetime in the RA message is 1800 seconds.

Configure the neighbor solicitation message retransmission interval

Command: `ipv6 nd base retrans-timer value`

View Mode: Configuration Mode

Default configuration: by default, the interface sends NS messages at an interval of 1000 milliseconds.

Configure the retransmission interval of the router in the RA message

Command: `ipv6 nd retrans-timer value`

View Mode: VLAN Interface Mode

Default configuration: by default, the value of the Retrans Timer field in the RA message issued by the interface is 0

Configure the time to keep a neighbor reachable

Command: `ipv6 nd base reachable-time value`

View Mode: Configuration Mode

Default configuration: by default, the interface remains reachable for 30000 milliseconds.

Configure the time to keep a neighbor reachable

Command: `ipv6 nd reachable-time value`

View Mode: VLAN Interface Mode

Default configuration: by default, the value of the Reachable Timer field in the RA message published by the interface is 0.

Configure the link MTU size

Command: `ipv6 nd link-mtu value`

View Mode: VLAN Interface Mode

Default configuration: by default, the value of the link MTU field in the RA message issued by the interface is 0.

When the source host sends a message from an interface, it compares the MTU of the interface with the Link MTU. If the message is longer than the smaller of the two values, that smaller value is used to fragment the message.

38.4.2 Configure the Number of Neighbor Solicitation Messages Sent for Duplicate Address Detection

After obtaining an IPv6 address, the interface sends a neighbor solicitation message for duplicate address detection. If it receives no response within the specified time, which is configured by the IPv6 nd retrans-timer command, the interface keeps sending the neighbor solicitation message. If no response is received after the number of transmissions reaches the configured number, the address is considered available.

Command	Description	CLI Mode
<code>ipv6 nd dad attempts <value></code>	By default, a neighbor solicitation message is sent 1 time during duplicate address detection. When the value is 0, duplicate address detection is disabled	Configuration Mode

38.5 IPv6 Static Route Configuration

Command	Description	CLI Mode
<code>ipv6 route <X:X::X:X/M> (<X:X::X:X> <ifName>) <distance></code>	Configure an IPv6 static route	Configuration Mode

38.6 IPv6 Display and Maintenance

After you complete the configuration above, run the show command in privileged view to display how IPv6 is operating, and check the displayed information to verify the effect of the configuration.

Command	Description	CLI Mode
<code>show ipv6 ndp nc</code>	Displays neighbor information	Privileged Mode
<code>show ipv6 interface (<ifName>) brief</code>	Displays IPv6 information for interfaces that can be configured with IPv6 addresses	Privileged Mode
<code>show ipv6 route (database)</code>	Displays IPv6 routes	Privileged Mode

Chapter 39 MLD SNOOPING Configuration

In a MAN or the Internet, when the same data packet is sent by unicast to multiple receivers in the network, the packet must be copied to each receiving endpoint. The number of packets to be sent therefore increases linearly with the number of receivers, which increases the overall burden on the host, the switching and routing equipment, and network bandwidth resources. Efficiency is greatly affected. With the increasing demand for multipoint video conferencing, video on demand, and group communication applications, multicast has increasingly become a popular transmission mode in multipoint communication, because it improves resource utilization.

The switch implements the MLD SNOOPING function to serve multicast applications. MLD SNOOPING monitors the MLD packets on the network to learn IPV6 multicast MAC addresses dynamically.

This chapter describes the concept and configuration of MLD SNOOPING. It covers the following topics:

- Introduction to MLD SNOOPING
- MLD SNOOPING configuration
- SNOOPING configuration example

39.1 Introduction to MLD SNOOPING

In a traditional network, multicast packets are treated as broadcasts within a subnet, which easily causes network traffic and network congestion. When MLD SNOOPING is implemented on the switch, it can dynamically learn IPV6 multicast MAC addresses and maintain the output port list for each IPV6 multicast MAC address, so the multicast data stream is sent only to those output ports. This reduces network traffic.

This section includes the following topics:

- MLD SNOOPING process
- Layer-2 dynamic multicast
- Join a group
- Leave a group

39.1.1 MLD SNOOPING Process

MLD SNOOPING is a Layer 2 network protocol. It monitors the MLD protocol packets that pass through the switch, maintains a multicast group according to the receiving port, VLAN ID, and multicast address of these MLD protocol packets, and then forwards the packets. Only the ports that have joined the multicast group can receive the multicast data stream, which reduces network traffic and saves network bandwidth.

The multicast group includes the multicast group address, member port, VLAN ID, and Age time.

An MLD SNOOPING multicast group is formed through a learning process. When a port of the switch receives an MLD REPORT packet, MLD SNOOPING creates a new multicast group and adds the port that received the MLD REPORT packet to it. When the switch receives an MLD QUERY packet, if the multicast group already exists on the switch, the port that received the MLD QUERY also joins the multicast group; otherwise the switch

simply forwards the MLD QUERY packet. MLD SNOOPING also supports the Done mechanism of MLD V2. If fast-leave is set to ENABLE for MLD SNOOPING, the receiving port can leave the multicast group immediately when it receives a Done packet of MLD V2. If fast-leave-timeout is configured, the multicast group waits for this time to expire before the port leaves the multicast group.

MLD SNOOPING has two update mechanisms. One is the Done mechanism described above. In most cases, MLD SNOOPING uses the age time to delete expired multicast groups. When a multicast group joins MLD SNOOPING, the joining time is recorded. When the multicast group has remained on the switch for longer than the configured age time, the switch deletes the multicast group.

When a port receives a Done protocol packet, the port is deleted immediately from the multicast group to which it belongs. This may affect the continuity of network data streams, because the port may be connected to a HUB or to a network device without the MLD SNOOPING function, and many devices that receive the multicast data stream may be connected to that device. When one device sends Done, it may affect the reception of the multicast data stream by the other devices. The Fast-leave-timeout mechanism can prevent this. The Fast-leave-timeout mechanism configures a leave waiting time. After receiving a leave packet, the port waits for the Fast-leave-timeout time before it is deleted from the multicast group to which it belongs. This can guarantee the continuity of the network multicast stream.

39.1.2 Layer-2 Dynamic Multicast

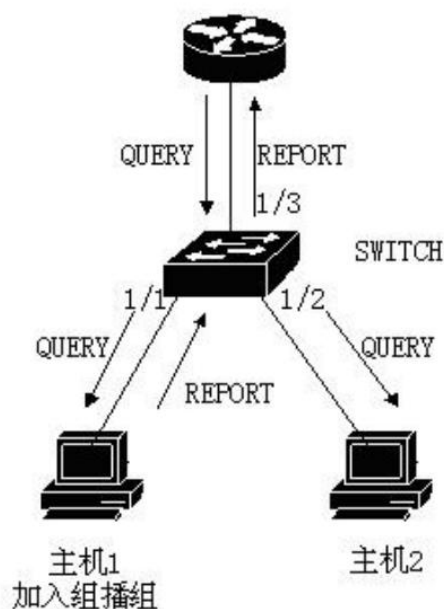
The multicast MAC address entries in the layer 2 hardware multicast forwarding table can be learned dynamically by MLD SNOOPING. The IPV6 multicast MAC address is learned dynamically through MLD SNOOPING.

When the switch disables MLD SNOOPING, the layer 2 hardware multicast forwarding table is in unregistered forwarding mode. Multicast MAC addresses cannot be learned dynamically, there is no entry in the layer 2 hardware multicast forwarding table, and all layer 2 multicast data streams are treated as broadcasts.

When the network has a multicast environment, you can enable MLD SNOOPING on the switch to control the multicast traffic of the network effectively. The layer 2 hardware multicast forwarding table is then in register forwarding mode, and the switch can learn multicast MAC addresses by monitoring the MLD protocol packets on the network. Only the layer 2 multicast flows that match the entries in the layer 2 hardware multicast forwarding table are forwarded.

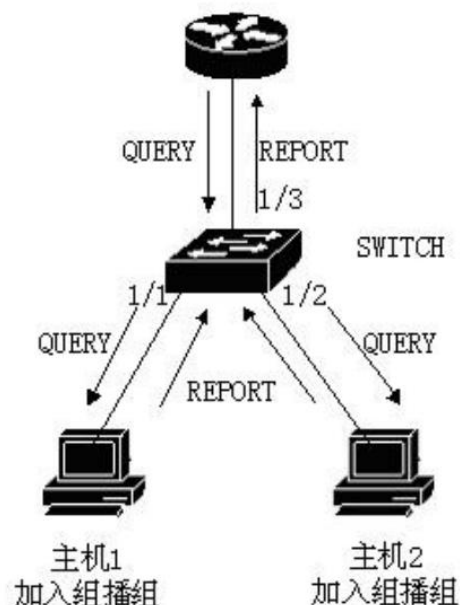
39.1.3 Join a Group

When a host wants to join a multicast group, it sends an MLD REPORT packet that specifies the multicast group it wants to join. When the switch receives an MLD QUERY packet, it forwards the packet to all other ports in the same VLAN. When a host under a port that wants to join the multicast group receives the MLD QUERY packet, it returns an MLD REPORT packet. When the switch receives the MLD REPORT packet, it creates a layer 2 multicast entry and adds the port that received the MLD QUERY packet and the port that received the MLD REPORT packet to that entry as its output ports.



As shown in the figure above, all the devices are in one subnet, and the VLAN of the subnet is assumed to be 2. The router runs the MLDv2 protocol and sends MLD QUERY packets periodically. Host 1 wants to join the multicast group ff15:1. When the switch receives the MLD QUERY packet on port 1/3, it records the port and forwards the packet to ports 1/1 and 1/2. Host 1 returns an MLD REPORT packet after receiving the MLD QUERY packet; host 2 does not send an MLD REPORT packet because it does not want to join the multicast group. When the switch receives the MLD REPORT packet on port 1/1, it forwards the packet out of query port 1/3 and creates a Layer 2 multicast entry, assuming the entry does not already exist, that includes the following:

Layer 2 Multicast Address	VLAN ID	List of Output Ports
33:33:00:00:00:01	2	1/1, 1/3



The conditions in the figure above are the same as in Figure 1. Host 1 has joined the multicast group ff15:1, and now host 2 wants to join the multicast group ff15:1. When host 2 receives the MLD query packet and returns an MLD REPORT packet, the switch receives the MLD REPORT on port 1/2, forwards the packet out of query port 1/3, and adds port 1/2 to the layer 2 multicast entry. The layer 2 multicast entry becomes:

Layer 2 Multicast Address	VLAN ID	List of Output Ports
33:33:5e:00:00:01	2	1/1, 1/2, 1/3

39.1.4 Leave a Group

To form a stable multicast environment, a device running MLD, such as a router, sends an MLD QUERY packet to all hosts at regular intervals. A host that has joined the multicast group or wants to join it receives the MLD QUERY and sends back an MLD REPORT.

A host can leave a multicast group in two ways: active departure and passive departure. Active leave means that the host sends an MLD leave packet to the router. Passive leave means that the host does not return an MLD REPORT after receiving an MLD QUERY from the router.

Corresponding to the way the host leaves the multicast group, a port on the switch can leave a layer 2 multicast entry in two ways: leaving on timeout and leaving on receipt of an MLD done packet.

When the switch does not receive an MLD report packet for a multicast group from a port within a certain time, the port is removed from the corresponding layer 2 multicast entry. If the layer 2 multicast entry then has no port, the entry is deleted.

When the fast-leave configuration of the switch is ENABLE and a port receives an MLD LEAVE packet for a multicast group, the port is removed from the corresponding layer 2 multicast entry. If the layer 2 multicast entry then has no port, the entry is deleted.

Fast-leave is generally used when a port is connected to a single host. If there is more than one host under a port, you can configure the fast-leave-timeout waiting time to ensure the continuity and reliability of the multicast stream in the network.

39.2 MLD SNOOPING Configuration

39.2.1 MLD SNOOPING Default Configuration

MLD SNOOPING is off by default, and the Layer 2 hardware multicast forwarding table is in unregistered forwarding mode.

Fast-leave is off by default.

Fast-leave-timeout is 300 seconds.

The default age time of the multicast group REPORT port is 400 seconds.

The default age time of the multicast group QUERY port is 300 seconds.

39.2.2 Opening and Closing MLD SNOOPING

You can turn the MLD SNOOPING protocol on globally or for individual VLANs. You can turn MLD SNOOPING on or off for a VLAN only when MLD SNOOPING is turned on globally.

Open global MLD SNOOPING

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping
```

Open MLD SNOOPING for a VLAN

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping vlan <vlan-id>
```

Turn off global MLD SNOOPING

```
Switch#configure terminal
Switch(config)#no ipv6 mld snooping
```

Turn off MLD SNOOPING for one VLAN

```
Switch#configure terminal
Switch(config)#no ipv6 mld snooping vlan <vlan-id>
```


39.2.3 Configure the Lifetime

Configure the lifetime of a multicast group

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping group-membership-timeout <interval> vlan <vlan-id>
```

Interval is in milliseconds.

Configure the lifetime of a query group

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping query-membership-timeout <interval> vlan <vlan-id>
```

Interval is in milliseconds.

39.2.4 Configure Fast-Leave

Start fast-leave for one VLAN

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping fast-leave vlan <vlan-id>
```

Close fast-leave

```
Switch#configure terminal
Switch(config)#no ipv6 mld snooping fast-leave vlan <vlan-id>
```

Configure the fast-leave wait times

```
Switch#configure terminal
Switch(config)# ipv6 mld snooping fast-leave-timeout <interval> vlan <vlan-id>
```

Restore the default fast-leave wait times

```
Switch#configure terminal
Switch(config)#no ipv6 mld snooping fast-leave-timeout vlan <vlan-id>
```

39.2.5 Configure MROUTER

Configure a static query port

```
Switch#configure terminal
Switch#interface ge1/6
Switch(config-ge1/6)#ipv6 mld snooping mrouter vlan [vlan-id]
```

39.2.6 Configure the Startup VLAN Querier

Enable the querier feature for VLAN 1

```
Switch#configure terminal
Switch(config)#ipv6 mld snooping querier vlan 1
```

39.2.7 Display Information

Display MLD SNOOPING configuration information

```
Switch#show ipv6 mld snooping
```

Display the configuration information of a VLAN

```
Switch#show ipv6 mld snooping vlan <vlan-id>
```

Display the aging information of a report multicast group

```
Switch#show ipv6 mld snooping age-table group-membership
```

Display the aging information of QUERY

```
Switch#show ipv6 mld snooping age-table query-membership
```

Display the forwarding information of a multicast group

```
Switch#show ipv6 mld snooping forwarding-table
```

Display MROUTER information

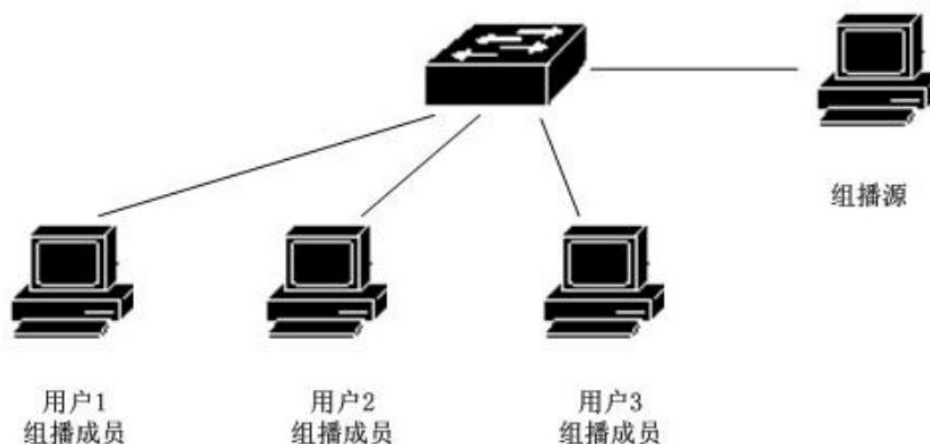
```
Switch#show ipv6 mld snooping mrouter
```

Display the current configuration of the system, including the MLD SNOOPING configuration

```
Switch#show running-config
```

39.3 MLD SNOOPING Configuration Example

Enable the MLD SNOOPING function on the switch. User 1, User 2, and User 3 can join a specific multicast group.



```
Switch#config t
Switch(config)#vlan database
Switch(config-vlan)#vlan 200
Switch(config)#ipv6 mld snooping
Switch(config)#interface ge1/1
Switch(config-ge1/1)#switchport mode access
Switch(config-ge1/1)#switchport access vlan 200
Switch(config)#interface ge1/2
Switch(config-ge1/2)#switchport mode access
Switch(config-ge1/2)#switchport access vlan 200
Switch(config)#interface ge1/3
Switch(config-ge1/3)#switchport mode access
Switch(config-ge1/3)#switchport access vlan 200
Switch(config)#ipv6 mld snooping group-membership-timeout 60000 vlan 200
```

Chapter 40 PoE Configuration

The switch supports the POE function. This chapter covers the following topics:

- Introduction to POE
- Configure POE

40.1 Introduction to POE

PoE (Power over Ethernet, also known as remote power supply) supplies power remotely to a device through an Ethernet interface, using a twisted pair to connect an external PD (Powered Device) such as an IP phone, a wireless AP, or a network camera.

(1) Benefits of PoE

- **Reliable:** centralized power supply, convenient backup
- **Simple and convenient connection:** the network terminal does not need an external power supply and needs only one network cable.
- **Standard:** it conforms to the IEEE 802.3af and 802.3at standards and uses the globally unified power interface.
- **Wide application prospects:** it can be used for IP phones, wireless APs (Access Points), portable device chargers, card readers, network cameras, and data acquisition.

(2) POE system composition

- A PoE system includes the PoE power supply, the PSE, and the PD.
- PoE power supply
- The PoE power supply supplies power to the whole PoE system. It is divided into two types: external power supply and built-in power supply.

PSE

- The PSE (Power Sourcing Equipment, Power Supply Equipment) is the board (daughter card). Each PSE independently manages the PoE interfaces on its board (daughter card). The PSE finds, detects, classifies, and supplies power to the PD on the line of the PoE interface. When it detects that the PD is unplugged, the PSE stops supplying power. Ethernet interfaces with PoE power supply are called PoE interfaces, and they include FE and GE interfaces.

PD

- A PD is a device that receives power from the PSE. A standard PD is a PD device that conforms to the IEEE 802.3af and 802.3at standards. While receiving the PoE power supply, the PD device is allowed to connect to other power supplies for power redundancy backup.

40.2 Configure POE

The POE configuration consists of the following three parts:

- Manual POE configuration
- POE policy configuration
- PD query configuration

40.2.1 Manual POE Configuration

The commands for manual POE configuration are as follows:

- Turn the POE power supply of the interface on or off.
- Display POE information

Command	Description	CLI Mode
[no] poe enable	Turn the POE power supply of the interface on or off. The power supply state of the interface is on by default.	Interface Configuration Mode
show poe	Displays POE information for all interfaces.	User Mode or Privileged User Mode

40.2.2 POE Policy Configuration

The commands for POE policy configuration are as follows:

- Open or close the POE policy of the interface
- Set the POE policy entry of the interface
- Display POE policy information

Command	Description	CLI Mode
[no] poe policy enable [no] poe policy shutdown clock <clock-value> week-day <day-value>	Turn the interface PoE policy on or off, and set or remove a PoE policy entry for the interface. The interface PoE policy is off by default, and no policy entry is set by default. The command can be entered more than once. clock-value is a time or time range in 24-hour format: 1 means 1 o'clock (between 1 and 2 o'clock), and 20-23 means 20 o'clock to 23 o'clock (between 20 and 0 o'clock). day-value is the day of the week, given as a single day or a run of consecutive days: 3 means Wednesday, and 1-7 means Monday to Sunday. A PoE policy takes effect only when the interface's PoE policy is turned on.	Interface Configuration Mode Interface Configuration Mode
show poe policy <if-name>	Displays POE policy information for an interface	User Mode or Privileged User Mode

40.2.3 PD Query Configuration

The command for the PD query configuration is as follows:

- Set the IP address of the PD
- Set the time interval for querying the PD
- Set the number of query timeouts for the PD
- Set the startup time of the PD
- Display PD information

Command	Description	CLI Mode
<code>poe pd-ip-address <ip-address> no poe pd-ip-address</code>	Sets or clears the IP address of the PD connected to the interface. By default, the IP address of the PD is not configured. If it is configured, the system queries the IP address periodically. If the PD does not respond for the given number of times, the system restarts the PD through the POE control.	Interface Configuration Mode
<code>poe pd-query-interval <interval> no poe pd-query-interval</code>	Sets the PD query interval. The default PD query interval is 5 seconds.	Interface Configuration Mode
<code>poe pd-timeout-number <number> no poe pd-timeout-number</code>	Sets the number of query timeouts for the PD. The default number of PD query timeouts is 3.	Interface Configuration Mode
<code>poe pd-boot-time <time> no poe pd-boot-time</code>	Sets the startup time of the PD. The default PD startup time is 120 seconds.	Interface Configuration Mode
<code>show poe pd-information</code>	Displays information for all configured PDs	User Mode or Privileged User Mode